Karl R. Popper

Post Scriptum a La lógica de la investigación científica. Vol. I

Realismo y el objetivo de la ciencia

Edición preparada por W. W. Bartley III

tecnos

SERIE DE FILOSOFIA Y ENSAYO

Dirigida por Manuel Garrido

Alston, W. P.; Edward, P.; Malcolm, N.; Nelson, J. O., y Prior, A.: *Los orígenes de la filosofía analítica. Moore, Russell, Wittgenstein.*
Austin, J. L.: *Sentido y percepción.*
Boden, M. A.: *Inteligencia artificial y hombre natural.*
Bottomore, T.; Harris, L.; Kiernan, V. G.; Miliband, R.; con la colaboración de Kolakowski, L.: *Diccionario del pensamiento marxista.*
Brown, H. I.: *La nueva filosofía de la ciencia.*
Bunge, M.: *El problema mente-cerebro.*
Couturat, Louis: *El álgebra de la lógica.*
Chisholm, R. M.: *Teoría del conocimiento.*
Díaz, Elías: *Revisión de Unamuno.* Análisis crítico de su pensamiento político.
Dopp, Joseph: *Nociones de lógica formal.*
Edelman, B.: *La práctica ideológica del Derecho.*
Fann, K. T.: *El concepto de filosofía en Wittgenstein.*
Ferrater Mora, J., y otros: *Filosofía y ciencia en el pensamiento español contemporáneo (1960-1970).*
Feyerabend, P.: *Tratado contra el método.* Esquema de una teoría anarquista del conocimiento.
García Suárez, A.: *La lógica de la experiencia.*
Garrido, M.: *Lógica simbólica.*
Gómez García, P.: *La antropología estructural de Claude Lévi-Strauss.*
Gurméndez, Carlos: *Ser para no ser.* Ensayo de una dialéctica subjetiva.
Hierro, José S.-P.: *Problemas de análisis del lenguaje moral.*
Hintikka, J.: *Lógica, juegos de lenguaje e información.* Temas kantianos de filosofía de la lógica.
Hintikka, J.: *Saber y creer.* Una introducción a la lógica de las dos nociones.
Kuhn, T. S.: *Segundos pensamientos sobre paradigmas.*
Lakatos, Imre, y otros: *Historia de la Ciencia y sus reconstrucciones racionales* (2.ª ed.).
Lindsay, P. H., y Norman, D. A.: *Introducción a la psicología cognitiva* (2.ª ed.).
Lorenzen, Paul: *Metamatemática.*
Lorenzo, J. de: *El método axiomático y sus creencias.*
Lorenzo, J. de: *La filosofía de la matemática de Jules Henri Poincaré.*
Martín Santos, Luis, y otros: *Ensayos de filosofía de la ciencia.* Simposio de Burgos. En torno a la obra de Karl R. Popper.
Mates, Benson: *Lógica matemática elemental.*
París, Carlos: *Hombre y naturaleza.*
Popper, Karl R.: *Búsqueda sin término.* Una autobiografía intelectual.
Popper, Karl R.: *Realismo y el objetivo de la ciencia.* Post Scriptum a la lógica de la investigación científica, vol. I.
Popper, Karl R.: *El universo abierto.* Post Scriptum a la lógica de la investigación científica, vol. II.
Popper, Karl R.: *Teoría cuántica y el cisma en física.* Post Scriptum a la lógica de la investigación científica, vol. III.
Prior, A. N.: *Historia de la lógica.*
Quine, W. V.: *La relatividad ontológica y otros ensayos.*
Quintanilla, Miguel A.: *Idealismo y filosofía de la ciencia.* Introducción a la epistemología de Karl R. Popper.
Rama, Carlos M.: *Teoría de la historia.* Introducción a los estudios históricos (3.ª ed.).
Rescher, Nicolás: *La primacía en la práctica.*
Robinet, André: *Mitología, filosofía y cibernética.*
Rodríguez Paniagua, José M.ª: *¿Derecho natural o axiología jurídica?*
Rodríguez Paniagua, José M.ª: *Marx y el problema de la ideología.*
Sahakian, W. S.: *Historia y sistemas de la psicología.*
Smart, J. J. C., y Williams, B.: *Utilitarismo: pro y contra.*
Sotelo, Ignacio: *Sartre y la razón dialéctica.*
Strawson, P. F.: *Ensayos lógico-lingüísticos.*
Vargas Machuca, R.: *El poder moral de la razón.* La filosofía de Gramsci.
Veldman, D. J.: *Programación de computadores en ciencias de la conducta.*
Wellman, C.: *Morales y éticas.*

REALISMO Y EL OBJETIVO DE LA CIENCIA

Digitized by the Internet Archive
in 2022 with funding from
Kahle/Austin Foundation

https://archive.org/details/postscriptumlalo0001popp

KARL R. POPPER

Post Scriptum a la lógica de la investigación científica

Volumen I

REALISMO Y EL OBJETIVO DE LA CIENCIA

Edición preparada por
W. W. Bartley III

tecnos

Título original:
Realism and the aim of science
From the Postscript to the Logic of Scientific Discovery
Edited by W. W. Bartley III

Traducción:
Marta Sansigre Vidal

Diseño de cubierta:
J. M. Domínguez y J. Sánchez Cuenca

Impresión de cubierta:
Gráficas Molina

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética o cualquier almacenamiento de información y sistema de recuperación, sin permiso escrito de Editorial Tecnos, S. A.

© Karl R. Popper, 1956, 1983
© EDITORIAL TECNOS, S. A., 1985
O'Donnell, 27 - 28009 Madrid
ISBN: 84-309-1228-2
Depósito Legal: M-42886-1985

Printed in Spain. Impreso en España por GAR. Villablino, 54. Fuenlabrada (Madrid)

A mi editor,
por la recuperación del Post Scriptum.

INDICE

PRINCIPALES PUBLICACIONES Y ABREVIATURAS DE TITULOS ..... *Pág.* 13
PROLOGO, por W. W. Bartley III 15
AGRADECIMIENTOS 21
INTRODUCCION DE 1982 23

REALISMO Y EL OBJETIVO DE LA CIENCIA

PREFACIO DE 1956 45

PARTE I

EL ENFOQUE CRITICO

CAPÍTULO I: LA INDUCCION 51
1 (* 1) Un filósofo desconcertado en el extranjero 52
2 (* 2) El enfoque crítico: la solución del problema de la inducción 58
3 (* 3) Sobre los llamados procedimientos inductivos, con notas sobre el aprendizaje y sobre el estilo inductivo 74
4 (* 4) Una familia de cuatro problemas de inducción 92
5 (* 5) Por qué la cuarta fase del problema es metafísica 111
6 (* 6) El problema metafísico 118
7 (* 7) Realismo metafísico 120
8 (* 8) La metafísica de Hume. Monismo «neutral» 128
9 (* 9) Por qué fracasa la teoría subjetivista del conocimiento .. 132
10 (* 10) Un mundo sin enigmas 143
11 (* 11) El status de las teorías y de los conceptos teóricos ... 147
12 (* 12) Crítica del instrumentalismo. El instrumentalismo y el problema de la inducción 152
13 (* 13) El instrumentalismo contra la ciencia 162
14 (* 14) La ciencia contra el instrumentalismo 168
15 (* 15) El objetivo de la ciencia 171

Nota sobre al numeración de las secciones. Las secciones de cada uno de los tres tomos del *Post Scriptum* y de las dos partes del primero están numeradas consecutivamente, empezando por la sección 1. Los números originales de las secciones, indicando el orden de éstas dentro de la totalidad del *Post Scriptum*, se dan entre paréntesis, precedidos de un asterisco, en los respectivos índices. Ed.

16 (* 16) Dificultades del realismo metafísico. Por un realista metafísico 187
CAPÍTULO II: LA DEMARCACION 199
17 (* 17) La significación del problema de demarcación 199
18 (* 18) Un caso de verificacionismo 202
19 (* 19) Contrastabilidad, pero no sentido 214
20 (* 20) Enunciados no contrastables 217
21 (* 21) El problema de «eliminar» la metafísica 219
22 (* 22) La asimetría entre «falsación» y «verificación» 221
23 (* 23) Por qué incluso las pseudo-ciencias pueden muy bien tener sentido. Programas metafísicos para la ciencia ... 229
CAPÍTULO III: METAFISICA: ¿CON SENTIDO O SIN SENTIDO? ... 234
24 (* 24) Observaciones lógicas sobre contrastabilidad y metafísica. 235
25 (* 25) Los términos metafísicos pueden definirse por medio de términos empíricos 251
26 (* 26) La cambiante filosofía del sentido y el sin sentido 254
CAPÍTULO IV: LA CORROBORACION 257
27 (* 27) Corroboración: certidumbre, incertidumbre, probabilidad. 257
28 (* 28) ¿«Corroboración» o «probabilidad»? 263
29 (* 29) ¿Corroboración o confirmación? 267
30 (* 30) El problema del grado de corroboración 270
31 (* 32) La corroboración 273
32 (* 32A) Algunos comentarios más sobre la definición del grado de corroboración 284
33 (* 33) El humanismo, la ciencia y el prejuicio inductivo 294
(* 31) *Addendum:* Observaciones críticas sobre el análisis de significados 301

PARTE II

LA INTERPRETACION PROPENSIVISTA DE LA PROBABILIDAD

CAPÍTULO I: PROBABILIDADES OBJETIVAS Y SUBJETIVAS 321
1 (* 34) Los sentidos de la probabilidad 321
2 (* 35) Probabilidades absolutas y relativas 323
3 (* 36) La interpretación propensivista. Interpretaciones objetivas y subjetivas 326
4 (* 37) Contrastaciones experimentales y su repetición: Independencia 328
5 (* 38) La interpretación lógica 331
6 (* 39) Comparación de la interpretación objetiva y la subjetiva. 334
7 (* 40) Las interpretaciones objetivistas y subjetivistas de «*b*» en «*p* (*a*,*b*)» 336
CAPÍTULO II. CRITICA DE LA INDUCCION PROBABILISTICA ... 341
8 (* 41) La regla inductiva simple 341
9 (* 42) Cómo interpretar la regla inductiva simple donde funciona 345
10 (* 43) Resumen del status «*b*» en «*p* (*a*,*b*)» 348
11 (* 44) La utilidad decreciente del aprendizaje por inducción ... 354
12 (* 45) La paradoja del aprendizaje inductivo 356

13 (* 46) Una máquina inductiva 359
14 (* 47) La imposibilidad de una lógica inductiva 363
15 (* 48) Lógica probabilística *versus* lógica inductiva 365
16 (* 49) La interpretación inductivista de la probabilidad 368
17 (* 50) La superfluidad de las teorías 372
18 (* 51) Carece de sentido contrastar una teoría 375
19 (* 52) Resumen de esta crítica 380
Addendum de 1981: Un breve resumen de la crítica de la inducción probabilística 382
CAPÍTULO III: OBSERVACIONES SOBRE LAS TEORIAS OBJETIVAS DE LA PROBABILIDAD 387
20 (* 53) El argumento en favor de las propensiones 388
21 (* 54) Donde tiene éxito la teoría frecuencial 400
22 (* 55) Donde fracasa la teoría frecuencial 407
23 (* 56) La significación del fracaso 412
24 (* 57) La teoría neoclásica y la frecuencial comparadas 413
25 (* 58) La estructura de la teoría neoclásica 422
26 (* 59) Enunciados de probabilidades singulares 424
27 (* 60) Otra crítica a las teorías subjetivas y lógicas 427
28 (* 61) Interpretación propensivista de la probabilidad de los sucesos singulares 430
Resumen final, 1982 434
INDICES 441

PRINCIPALES PUBLICACIONES Y ABREVIATURAS DE TITULOS

En las *Notas* se han utilizado las siguientes abreviaturas para referirse a las principales publicaciones del autor:

L. d. F.:	*Logik der Forschung,* Julius Spriner Verlag, Viena, 1934.
O. S.:	*The Open Society and Its Enemies,* Princeton University Press, 1950. Traducida como *La sociedad abierta y sus enemigos,* Editorial Paidós, Buenos Aires, 1957.
The Poverty:	*The Poverty of Historicism* (1944-45), 1957, 1960. Traducida como *La miseria del Historicismo,* Taurus Ediciones, Madrid, 1961.
L. Sc. D.:	*The Logic of Scientific Discovery,* Hutchinson & Co., Londres; Basic Books, Nueva York, 1959. Traducida como *La lógica de la investigación científica,* Editorial Tecnos, Madrid, 1982.
C. & R.:	*Conjetures and Refutations: The Growth of Scientific Knowledge,* Routledge and Kegan Paul, Londres; Basic Books, Nueva York, 1963. Traducida como *El desarrollo del conocimiento científico: Conjeturas y Refutaciones,* Editorial Paidós, Buenos Aires, 1967.
Obj. Kn.:	*Objective Knowledge: An Evolutionary Approach,* Clarendon Press, Oxford, 1972. Traducida como *Conocimiento objetivo: Un enfoque evolucionista,* Editorial Tecnos, Madrid, 1982, 2.ª ed.
U. Q.:	*Unended Quest: An Intellectual Autobiography,* Fontana-Collins, Londres, 1976. Traducida como *Búsqueda sin término: Una autobiografía intelectual,* Editorial Tecnos, Madrid, 1977.
S. & I. B.:	*The Self and Its Brains,* en colaboración con Sir John Eccles, 1977. Traducida como *El yo y su cerebro,* Editorial Labor, Barcelona, 1980.
O. U.:	*The Open Universe: An Argument for Indeterminism.* Vol. II of *The Postscript to the Logic of Scientific Discovery,* Hutchinson, Londres. Traducida como *El universo abierto: Un argumento en favor del indeterminismo,* Editorial Tecnos, Madrid, 1984.
Q. T. & S. Ph.:	*Quantum Theory and the Schism in Physics,* Vol. III of *The Postscript to the Logic of Scientific Discovery,* Hutchinson, Londres, 1982. Traducida como *Teoría cuántica y el cisma en física,* Editorial Tecnos, Madrid, 1985.

PROLOGO

Este libro, Realismo y el objetivo de la ciencia, *es el primer tomo del tan esperado* Post Scriptum a La lógica de la investigación científica *de Sir Karl Popper. Aunque se escribió hace unos veinticinco años, no se había publicado todavía. Contiene un nuevo desarrollo, sumamente ampliado, de las concepciones de Popper, sobre la inducción, la demarcación y la corroboración y presenta también su teoría propensivista de la probabilidad. Este libro contiene asimismo una consideración detallada de las críticas y objeciones que se han hecho a las concepciones de Popper a lo largo de los años transcurridos desde la publicación inglesa de* La lógica de las investigación científica, *así como las réplicas a dichas críticas y objeciones.*

Este tomo se escribió, junto con las demás partes del Post Scriptum *(todas las cuales se publican ahora), principalmente durante los años de 1951 a 1956, en un momento en que* Logik der Forschung, *el primer libro publicado de Popper (1934), estaba siendo traducido al inglés con el título de* The Logic of Scientific Discovery.

Los distintos volúmenes del Post Scriptum *formaban parte, originariamente, de una serie de Apéndices a* La lógica de la investigación científica, *en la que Popper se proponía corregir, extender y desarrollar las ideas de su primer libro. Algunos de esos Apéndices se incluyeron finalmente en* La lógica de la investigación científica, *cuando se publicó en 1959. Pero un grupo de Apéndices adquirió vida propia y se fue convirtiendo, gradualmente, en un trabajo bien integrado y distinto que excedía en longitud al original* Logik der Forschung. *Se decidió publicar esta nueva obra —titulada* Postscript: After Twenty Years [*Post Scriptum: Veinte años después*]— *como corolario o volumen compañero de* La lógica de la investigación científica *y, en consecuencia, se imprimió en galeradas en 1956-1957.*

Pocos meses antes de la fecha prevista para la publicación, sin embargo, el proyecto se paró en seco. En Búsqueda sin término, *su autobiografía intelectual, Sir Karl dice de esas galeradas: «La corrección de pruebas se convirtió en una pesadilla... Entonces tuve que someterme a una operación en ambos ojos. Después de ella no pude entregarme de nuevo a la corrección de pruebas hasta pasado algún tiempo y, como resultado, el* Post Scriptum *tuvo que ser* [illegible]

Recuerdo claramente aquellos momentos. Fui a Viena a visitar a Popper al hospital poco después de su operación por varios desprendimientos en las dos retinas; y trabajamos sobre el Post Scriptum *mientras él se iba recuperando. Durante mucho tiempo apenas veía y temíamos que se quedase ciego.*

Cuando recobró la vista trabajó mucho en el Post Scriptum: *añadió varias secciones e hizo miles de correcciones a las galeradas. Pero la presión del resto de su trabajo se hacía excesiva y después de 1962 apenas se añadió nada al texto. Durante la década siguiente, enormemente productiva, Popper publicó* El desarrollo del conocimiento científico: Conjeturas y refutaciones *(1963), y completó y publicó tres libros nuevos:* Conocimiento objetivo: Un enfoque evolucionista *(1972),* Búsqueda sin término *(1974 y 1976) y (con Sir John Eccles)* El yo y su cerebro *(1977), así como numerosos artículos. Esos fueron los años, y las obras, en los que su ahora famosa teoría de la mente objetiva (y de los Mundos, 1, 2 y 3) se desarrolló y en los que su enfoque se extendió a las ciencias biológicas.*

Mientras tanto, el Post Scriptum, *que representaba la culminación del trabajo de Sir Karl en la filosofía de la física, seguía sin publicarse. Pero no sin leerse: la mayor parte de los alumnos y colegas más allegados de Popper han estudiado esta obra y, a lo largo de los años, varios han poseído copias de las galeradas. Para los que, como yo, han conocido este libro y han sido profundamente influidos por él, es una gran satisfacción verlo finalmente completo y compartirlo con el gran público.*

El texto que se ha preparado ahora para su publicación es, esencialmente, el que existía en 1962. Excepto en algunos lugares, que se han señalado, no se han hecho alteraciones importantes. Pensamos que éste era el enfoque apropiado para una obra que ha adquirido, a través de su influencia en los alumnos y los colegas de Popper, un carácter histórico —al haber transcurrido unos veinticinco años desde que se redactó y unos cuarenta y cinco desde que se escribió la Logik der Forschung *original—. Es obvio que muchos puntos se hubieran expresado hoy de otra forma. Pero una revisión completa por parte del autor habría retrasado indefinidamente la publicación del libro.*

La preparación para la publicación ha consistido en reunir diferentes versiones de algunas partes del texto que se habían ido acumulando a lo largo de los años; en revisar el libro y en añadir notas bibliográficas y de otra índole para ayudar al lector. Las pocas nuevas adiciones del propio Popper aparence con asterisco. Mis breves notas, bibliográficas y editoriales, aparecen también entre corchetes, seguidas de la abreviatura «Ed.». Aquí he seguido, en general, la práctica establecida por Troels Eggers Hansen, editor de Die beiden Grundprobleme der Erkenntnis-

theorie *de Popper (escrita en 1930-1932 y publicada en 1979). Popper ha podido comprobar el trabajo editorial en una serie de reuniones que hemos tenido en distintos lugares en los dos años últimos —en Heidelberg, Guelph, Toronto, Washington DC, Schloss Kronberg y en su casa de Buckinghamshire. Popper ha añadido nuevos prefacios a todos los volúmenes y un nuevo epílogo al segundo.*

Se ha hecho un cambio importante en la presentación por sugerencia mía. Publicar esta extensa obra en un solo tomo hubiera sido posible, pero hubiera resultado un libro pesado y difícil de manejar, inasequible para muchos estudiantes de filosofía. Algunas partes del Post Scriptum *tienen un amplio interés, no sólo para los filósofos y estudiantes de filosofía, sino también para aquellas personas, de entre el gran público, que estén interesadas por la ciencia contemporánea y su interpretación, así como por la cosmología.*

Estas partes son también, en su totalidad, independientes unas de otras. Esto me llevó a sugerir que se publicase la obra en tres tomos independientes del mismo formato, que en conjunto integrasen el Post Scriptum. *Después de algunas dudas, Sir Karl aceptó la propuesta, y también los títulos que yo le sugerí para los tres volúmenes.*

El Post Scriptum *se publica, pues, como sigue:*

Realismo y el objetivo de la ciencia *(tomo I);*

El universo abierto: Un argumento en favor del indeterminismo *(tomo II);*

Teoría cuántica y el cisma en física *(tomo III).*

Aunque los tres volúmenes del Post Scriptum *pueden leerse por separado, el lector debe saber que forman un argumento unitario. Cada volumen del* Post Scriptum *ataca uno u otro de los enfoques subjetivistas o idealistas del conocimiento; cada uno estructura un componente o más para un enfoque objetivo y realista del conocimiento.*

Así, en este tomo, Popper acosa al «inductivismo», al que considera como principal fuente del subjetivismo e idealismo, en cuatro planos: lógico, metodológico, epistemológico y metafísico. Popper desarrolla su teoría de la falsabilidad y traza sus efectos, deslindando las opiniones científicas de las no científicas y de las pseudo-científicas. Y presenta su teoría sobre la corroboración como una forma de expresar preferencia racional por una teoría frente a otra sin recurrir a las «certezas» subjetivas o a la «justificación» de las filosofías convencionales. En el primer volumen Popper trata también su relación con los personajes históricos de la filosofía, tales como Berkeley, Hume, Kant, Mach y Russell, que han hecho contribuciones importantes a la tradición subjetivista; y da respuestas detalladas a críticos filósofos y científicos

contemporáneos. En este volumen Popper ataca también la interpretación subjetiva del cálculo de probabilidades, una interpretación que tiene sus raíces en la creencia de que la probabilidad mide un estado subjetivo de conocimiento insuficiente. En La lógica de la investigación científica, *Popper defendía una interpretación objetiva del cálculo de probabilidades, para la que utilizaba la interpretación frecuencial. Ahora Popper critica también la interpretación frecuencial. En su lugar presenta con todo detalle su propia interpretación propensivista —una interpretación que, durante los últimos veinte años, ha encontrado muchos defensores—. Estas ideas y estos argumentos se aplican y se desarrollan en los volúmenes restantes.*

En El universo abierto: Un argumento en favor del indeterminismo *Popper desarrolla su propia perspectiva indeterminista y presenta la crítica tanto de la forma «científica» como de la forma metafísica del determinismo y arguye que la física clásica, al contrario de lo que normalmente se cree, no presupone ni implica determinismo más de lo que lo implica la física cuántica. Y, sin embargo, encuentra que el determinismo metafísico continúa subyacente en la obra de muchos teóricos cuánticos contemporáneos, incluidos los oponentes del determinismo. Popper encuentra el origen del papel que siguen jugando dentro de la física las interpretaciones subjetivas de la probabilidad en esas presuposiciones metafísicas deterministas.*

Hay una profunda conexión entre los argumentos de los tomos primero y segundo, en la preocupación que muestran por la libertad, la creatividad y la racionalidad del hombre.

El primer tomo, en su consideración de la justificación y de la racionalidad, rebate una pretensión subjetivista y escéptica sobre los límites de la crítica —y con ello, sobre los límites de la racionalidad—. Si existiera ese límite, sería inútil el debate serio; y su apariencia sería ilusoria.

El segundo tomo, en su tratamiento del determinismo, defiende la idea de que nuestra racionalidad está limitada con respecto a la predicción del aumento futuro del conocimiento humano. Si tal límite no *existiera, entonces sería inútil el debate serio; y su apariencia sería ilusoria.*

Popper mantiene, pues, que la razón humana no tiene límites con respecto a la crítica y, sin embargo, los tiene con respecto a su capacidad de predicción; y demuestra que la falta de limitación en un caso, y la limitación en el otro, son absolutamente necesarias para la existencia de la racionalidad humana.

En el tomo tercero, Teoría cuántica y el cisma en física, *Popper reexamina y rebate una serie de argumentos y «paradojas» extensamente utilizados para defender un punto de vista idealista. Popper conjetura que los problemas de interpretación de la*

mecánica cuántica pueden reducirse a problemas de interpretación del cálculo de probabilidades, y a partir de ahí, desarrolla con más amplitud su propia interpretación propensivista de la probabilidad. Y a continuación expone su extensa crítica de algunas de las interpretaciones dominantes de la teoría cuántica, tratando de resolver sus conocidas paradojas y de exorcizar al «observador» de la física cuántica.

En su concluyente «Epílogo metafísico», Popper teje los temas de todo el Post Scriptum *en un estudio histórico y programático del papel de los programas metafísicos de investigación o de las interpretaciones en la historia de la física.*

El editor quiere expresar su gratitud a The American Council of Learned Societies *y a* The American Philosophical Society *por el generoso apoyo prestado a su trabajo editorial en estos volúmenes; y a los Profesores Donald T. Campbell y F. A. von Hayek por sus consejos y su apoyo. También desea agradecer a su secretaria, Nancy Artis Sadoyama, su dedicación e incansable ayuda.*

W. W. Bartley III

AGRADECIMIENTOS

Quiero aprovechar esta oportunidad para agradecer a mi colega John W. N. Watkins el ánimo que me ha dado con su constante interés. Ha leído este volumen en manuscrito y en galerada y ha hecho sugerencias muy útiles para mejorarlo. Fue por sugerencia suya por lo que decidí publicar el *Post Scriptum* como una obra aparte, en lugar de como había pensado originariamente, como una serie de Apéndices a *La lógica de la investigación científica*. Pero, aún más importante para la terminación de la obra, fue el interés que mostró por las ideas que contiene.

Quiero también dar las gracias a los cotraductores de *The Logic of Scientific Discovery*, Dr. Julius Freed y Lan Freed, que leyeron la mayor parte de este volumen en galeradas e hicieron numerosas sugerencias para mejorar el estilo. [Los dos murieron muchos años antes de que se publicara el libro. *Ed.*]

Joseph Agassi fue, durante el tiempo en que se escribió este libro, primero estudiante de investigación y luego mi ayudante de investigación. Con él discutí casi cada sección, con todo detalle, y con frecuencia, como consecuencia de sus consejos, desarrollé uno o dos enunciados hasta formar una nueva sección completa o, en su caso, una nueva parte [la Parte II de *Realismo y el objetivo de la ciencia*]. Su cooperación fue para mí de enorme valor.

Quiero dar las gracias también a la London School of Economics and Political Science, que me hizo posible la provechosa ayuda del Dr. Agassi, y al Center for Advanced Study in the Behavioral Sciencies (Ford Foundation) en Standfort, California, porque me dio la oportunidad de trabajar ininterrumpidamente, desde octubre de 1956 a julio de 1957, en las galeradas de este libro, y porque hizo posible que el Dr. Agassi me ayudara durante ese tiempo.

PENN, BUCKINGHAMSHIRE, 1959

El Profesor W. W. Bartley III, fue alumno mío y, más tarde, colega en la London School of Economics, de 1958 a 1963, y colaboró estrechamente conmigo en 1960-1962 en este libro. En 1978 tuvo la amabilidad de consentir en actuar como editor del *Post Scriptum*. Le estoy muy agradecido por su asistencia y por emprender tan ardua tarea. Le debo más de lo que puedo expresar.

Es también un placer dar las gracias a varias personas más, que en los años pasados, trabajaron conmigo en este *Post Scriptum*, en particular, Alan E. Musgrave, David Miller, Arne F. Petersen, Tom Settle y Jeremy Shearmur. De ellos, David Miller y Arne Petersen deben ser mencionados especialmente a causa de la enorme cantidad de trabajo que ambos realizaron en varios períodos antes de 1970.

La London School ha continuado ayudándome durante todos estos años, designando un ayudante de investigación. Durante los trece años transcurridos desde mi jubilación, en 1969, lo ha hecho con ayuda de una beca de la Nuffield Foundation, a la que quiero expresar mi agradecimiento. Los principales responsables de esta disposición fueron mi amigo y sucesor, el Profesor John Watkins; el difunto Sir Walter Adams, director de la London School, y el director actual, profesor Ralf Dahrendorf, por cuya entrañable amistad e interés por mi trabajo estoy profundamente agradecido.

Si el *Post Scriptum* se hubiera publicado en la década de 1950, se lo hubiera dedicado a Bertrand Russell: El Profesor Bartley me ha dicho que hay una carta que lo atestigua en los Archivos de Ruseel en McMaster University.

Debo mencionar, finalmente, que el *Post Scriptum* (junto con la traducción inglesa de la *Lógica de la investigación científica*) me parecía casi terminado en 1954. Fue entonces cuando escogí su título original *Postscript: After Twenty Years* [Post Scriptum: Veinte años después], aludiendo a la publicación de *Logik der Forschung* en 1934.

PENN, BUCKINGHAMSHIRE, 1982

INTRODUCCION DE 1982

En esta introducción al primer tomo del *Post Scriptum* quisiera discutir brevemente varias cuestiones que se han planteado durante las tres décadas transcurridas desde que se escribió este libro, contra las concepciones que en él presento.

I

La primera cuestión se refiere a los términos técnicos «falsable» (empíricamente refutable) y «falsabilidad» (refutabilidad empírica). Introduje estos términos por primera vez en *Erkenntnis* 3, 1933, y en *Logik der Forschung,* 1934, en relación con mi solución al problema de la demarcación (que se discute extensamente en la Parte I, capítulo II, de este tomo). El problema de la demarcación consiste en encontrar un criterio que nos permita distinguir entre enunciados que pertenezcan a las ciencias empíricas (teorías, hipótesis) y otros enunciados, en particular los enunciados pseudo-científicos, pre-científicos y metafísicos; pero también los enunciados lógicos y matemáticos. El problema de la demarcación debe distinguirse del problema, mucho más importante, de la verdad: teorías de las que se ha demostrado que eran falsas —como, por ejemplo, las fórmulas de radiación de Rayleig-Jeans y de Wien, o el modelo atómico de Bohr de 1913— pueden, no obstante, retener el carácter de hipótesis científicas, empíricas.

Aunque, siguiendo a Tarski, no creo que sea posible un criterio de verdad, he propuesto un criterio de demarcación —el criterio de falsabilidad—. Mi propuesta es que un enunciado (una teoría, una conjetura) cumple las condiciones para pertenecer a las ciencias empíricas si, y sólo si, es falsable.

Pero, ¿cuándo es falsable un enunciado? Es de gran importancia para la presente discusión advertir que la falsabilidad en el sentido de un criterio de demarcación es un asunto puramente lógico. Se refiere únicamente a la estructura lógica de los enunciados y de las clases de enunciados. Y *no* tiene *nada* que ver con la cuestión de si ciertos resultados experimentales posibles podrían o no ser aceptados como falsaciones.

Un enunciado o teoría es falsable, según mi criterio, si y sólo si, existe como mínimo un falsador potencial —al menos un

enunciado básico que esté en conflicto con él—. Es importante no exigir que el enunciado básico en cuentión sea *verdadero*. La clase de enunciados básicos está concebida de modo que un enunciado básico describa un suceso lógicamente posible que sea lógicamente posible que pueda ser observado.

Para hacer esto menos abastracto daré aquí cuatro ejemplos: dos de enunciados falsables y dos de enunciados infalsables.

1) «Todos los cisnes son blancos.» Esta teoría es falsable puesto que, por ejemplo, contradice el enunciado básico siguiente (que es, por cierto, falso): «El 16 de mayo de 1934, entre las diez y las once de la mañana, hubo un cisne negro parado ante la estatuta de la Emperatriz Isabel en el Volksgarten de Viena».

2) El principio de Einstein de la proporcionalidad entre la masa inerte y la masa (pasivamente) pesada. Este principio de equivalencia entra en conflicto con muchos falsadores potenciales: sucesos cuya observación es lógicamente posible. Sin embargo, a pesar de todos los intentos (los experimentos de Eötvös, recientemente refinados por Dicke) de llevar a cabo experimentalmente esa falsación los experimentos han corroborado, hasta el momento, el principio de equivalencia.

3) «Todas las acciones humanas son egoístas, motivadas por el propio interés.» Esta teoría es ampliamente aceptada: tiene variantes en el conductismo, el psicoanálisis, la psicología individual, el utilitarismo, el marxismo vulgar, la religión y la sociología del conocimiento. Está claro que esta doctrina, con todas sus variantes, no es falsable: ningún ejemplo de acción altruista puede refutar la concepción de que había un motivo egoísta oculto tras ella.

4) Los enunciados puramente existenciales no son falsables, como el famoso ejemplo de Rudolf Carnap: «Existe un color ('rojo-eupatorio') que produce terror a quienes lo miran». Otro ejemplo es: «Hay una ceremonia cuya realización exacta hace aparecer al diablo». Tales enunciados no son falsables. (Son, en principio, verificables: es lógicamente posible encontrar una ceremonia cuya realización lleve a la aparición de una figura humana con cuernos y pezuñas. Y si la repetición de la ceremonia no consigue el mismo resultado, eso no supone falsación, porque quizá se haya omitido un aspecto inadvertido, pero esencial de la ceremonia correcta.)

Como muestran esos ejemplos, la falsabilidad en el sentido del criterio de demarcación no significa que pueda realizarse en la práctica una falsación o que, si se lleva a cabo, carezca de problemas. La falsabilidad en el sentido del criterio de demarcación no significa nada más que una relación lógica entre la

teoría en cuestión y la clase de los enunciados básicos o la clase de los sucesos descritos por ellos: los falsadores potenciales. La falsabilidad es, pues, relativa a esas dos clases: si una de estas clases está dada, la falsabilidad es una cuestión de pura lógica: del carácter lógico de la teoría bajo consideración.

Que es necesario que nos den la clase de falsadores potenciales (o de enunciados básicos) puede mostrarse fácilmente con nuestro primer ejemplo: «todos los cisnes son blancos».

Como ya he dicho, este enunciado es falsable. Supongamos, sin embargo, que hay alguien que, cuando se le enseña un cisne que no es blanco, adopta la posición de que no puede ser un cisne, ya que es «esencial» que un cisne sea blanco.

Tal posición equivale sostener que los cisnes no blancos son estructuras lógicamente imposibles (y, por tanto, inobservables). Los excluye como clase de falsadores potenciales.

En relación con esta clase modificada de falsadores potenciales, el enunciado «todos los cisnes son blancos» es, naturalmente, infalsable. Para evitar ese paso podemos exigir que cualquiera que defienda el carácter científico-empírico de una teoría sea capaz de especificar bajo qué condiciones estaría dispuesto a considerarla falsada. es decir, tendría que ser capaz de describir al menos algunos falsadores potenciales.

Llegamos ahora al segundo sentido de «falsable» y «falsabilidad», que debe distinguirse muy claramente de mi criterio puramente lógico de demarcación para evitar confusiones burdas.

Uno puede plantear la cuestión de si la falsación efectiva es alguna vez tan clara que *haya que* considerar falsada la teoría en cuestión (y haya, por tanto, que considerarla falsa). ¿No hay siempre una salida para alguien que quiera salvar la teoría en cuestión?

Siempre he mantenido, incluso en la primera edición de *Logik der Forschung* (1934), y también en mi libro anterior, aunque recientemente publicado, *Die beiden Grudprobleme der Erkenntnistheorie* (1979, escrito entre 1930 y 1933), que nunca es posible demostrar concluyentemente que una teoría científica empírica es falsa. En *este sentido,* esas teorías *no son falsables.* «Todo sistema teórico puede ser protegido de diversas maneras contra una falsación empírica» (*Grundprobleme,* p. 353). «Siempre es posible encontrar alguna vía de escape de la falsación, por ejemplo mediante la introducción *ad hoc* de una hipótesis auxiliar...» (*Lógica de la investigación científica* [*L. I. C.*], p. 41, en la misma sección en que se introduce la falsabilidad). «No es posible jamás presentar una refutación concluyente de una teoría...» (*L. I. C.* p. 49).

De ahí, repito, que tengamos que distinguir entre dos significados de las expresiones «falsable» y «falsabilidad»:

1) «Falsable», como término lógico-técnico, en el sentido del criterio de demarcación por la falsabilidad. Este concepto puramente lógico —falsable en principio, podríamos decir— descansa en una relación lógica entre la teoría en cuestión y la clase de enunciados básicos (o los falsadores potenciales por ellos descritos).

2) «Falsable» en el sentido de que la teoría en cuentión puede se falsada *definitiva o concluyente o demostrablemente* («demostrablemente falsable»). Siempre he insistido que, incluso una teoría que es obviamente falsable en el primer sentido, no lo es nunca en el segundo. (Por esta razón he usado la expresión «falsable», por norma general, sólo en el primer sentido, el sentido técnico. En el segundo sentido he hablado, por norma general, no de «falsabilidad», sino más bien de la «falsación» y de sus problemas.)

Está claro que los sufijos «-able» y «-abilidad» están usados de forma algo diferente en esos dos sentidos. Aunque el primer sentido se refiere a la posibilidad lógica de una falsación en principio, el segundo caso se refiere a una *prueba experimental práctica y concluyente* de flasedad. Pero no existe nada parecido a una prueba concluyente para resolver una cuestión empírica.

Toda una literatura descansa sobre el hecho de que no se haya observado esta distinción. Se dice a menudo que mi criterio de demarcación es inaplicable porque las teorías científicas empíricas no pueden falsarse definitivamente. Aunque es menos importante, también se dice (véase sección IV, más adelante) que el descubrimiento de la infalsabilidad de las teorías científicas en el segundo sentido es un logro que contradice mi teoría, a pesar de que he sido yo mismo quien ha señalado ese hecho una y otra vez. (En lugar de diferenciar los dos significados —«falsabilidad$_1$», la posibilidad de que ciertas teorías puedan falsarse en principio, porque tienen algunos falsadores potenciales, y «falsabilidad$_2$», la *siempre* problemática posibilidad de que pueda mostrarse que una teoría es falsa, puesto que no existen pruebas empíricas definitivas— se han hecho distinciones irónicas entre «Popper$_0$», «Popper$_1$», «Popper$_2$» y así sucesivamente [es decir, entre varias etapas de «Popper» que se contradicen notoriamente unas a otras y no pueden armonizarse][1]. Y las dificultades, en muchos casos la imposibilidad, de

[1] [Véanse las obras de Imre Lakatos, especialmente «Criticism and the Methodology of Scientific Research Programmes» *Proc. Arist. Soc.* 69, pp. 149-186, y *The Methodology of Scientific Research Programmes*, 1978, pp. 93-101. *Ed.*]

una falsación práctica concluyente se presentan como una dificultad o incluso la imposibilidad del criterio de demarcación propuesto.)

Todo esto tendría poca importancia si no fuera por el hecho de que ha llevado a algunas personas a abandonar el racionalismo en la teoría de la ciencia y a lanzarse al irracionalismo. Porque si la ciencia no avanza racional y críticamente, ¿cómo podemos tener la esperanza de que se tomen decisiones racionales en otros terrenos? Un ataque, hecho a la ligera, a un término lógico-técnico mal entendido ha llevado, pues, a algunos a conclusiones filosóficas, e incluso políticas, desastrosas y de largo alcance.

Debería subrayarse que la incertidumbre de toda falsación empírica (que yo he señalado repetidas veces) no debe tomarse demasiado en serio (como también he señalado). Hay una serie de falsaciones importantes que son tan «definitivas» como lo permite la general falibilidad humana. Además, toda falsación puede, a su vez, volver a ser contrastada. Un ejemplo de una falsación —la falsación del modelo del átomo de Thomson, que llevó a Ernest Rutherford a proponer el modelo nuclear— debe mencionarse aquí para ilustrar la fuerza que puede tener una falsación. En el modelo del átomo de Thomson, la carga positiva se distribuía por todo el espacio ocupado por el átomo. Rutherford había aceptado ese modelo. Pero entonces vinieron los famosos experimentos de sus alumnos, Geiger y Marsden. Encontraron que las partículas alfa que se disparaban sobre un trozo de pan de oro muy fino, a veces eran reflejadas desde el pan de oro, en lugar de ser sólo desviadas. Las partículas reflejadas eran escasas —aproximadamente una entre veinte mil—, pero aparecían con regularidad estadística. Rutherford estaba atónito. Un cuarto de siglo después escribió sobre ello. «Fue realmente el acontecimiento más increíble que me ha ocurrido en toda mi vida. Era casi tan increíble como si disparases un proyectil de quince pulgadas contra un trozo de papel de seda y rebotase y te diese»[2].

La formulación de Rutherford es excelente. No es imposible —ciertamente no es *lógicamente* imposible— que un disparo de un cañón gigante sobre un trozo de papel de seda sea reflejado por éste, incluso con una probabilidad estadística regular de 1 sobre 20.000. Esto no es lógicamente imposible; y de ahí que la teoría de Thomson (según la cual los átomos forman un muro como el papel de seda) no esté refutada definitivamente. Pero

[2] Lord Rutherford, «The Development of the Theory of Atomic Structure», en J. Needham y W. Pagel, eds., *Background of Modern Science*, 1938, p. 68.

Rutherford y otros físicos, entre ellos Niels Bohr, llegaron a la conclusión de que hacía falta otra teoría. Por tanto, propusieron que la teoría de Thomson se considerase falsada y que se sustituyese por el modelo nuclear de Rutherford; y algo más tarde (al presentar dicho modelo sus propios problemas) por el excelente modelo atómico de Bhor que, al cabo de unos doce años, fue a su vez sustituido por la mecánica cuántica.

Con frecuencia se tarda bastante tiempo en aceptar una falsación. Normalmente no se acepta hasta que la teoría falsada es sustituida por la propuesta de una teoría nueva y mejor. Como advirtió Max Planck, a veces hay que esperar a que se haya formado una nueva generación de científicos; eso, sin embargo, no es siempre necesario. No lo fue con el nuevo modelo del átomo de Rutherford (1912) ni con el reconocimiento de J. J. Thomson (1897) de partículas subatómicas tales como el electrón, que significaban la falsación de la teoría del átomo indivisible. (Los átomos se habían considerado indivisibles por definición desde el año 460 a. C. aproximadamente.) Tampoco fue así con la falsación de Carl Anderson (1932) de la poderosa teoría de que sólo había dos partículas elementales —el protón y el electrón— ni con el rechazo de Hideki Yukawa de la teoría electro-magnética de la materia.

Estos son sólo cuatro entre muchos ejemplos de revoluciones científicas que se introdujeron gracias a falsaciones de éxito.

El significado lógico-técnico, mal entendido, de falsabilidad en el primer sentido, en el sentido del criterio de demarcación, ha llevado a dos leyendas históricas. La primera leyenda, poco importante, es que yo pasé por alto el hecho de que la falsabilidad de las teorías no es concluyente —el hecho de que las teorías no son concluyentemente falsables en el segundo sentido—. Cuando, en realidad, yo lo había resaltado repetidamente desde 1932. La segunda leyenda (y ésta es mucho más importante) es que la falsación no juega ningún papel en la historia de la ciencia. En realidad, juega un papel primordial, a pesar de su carácter no definitivo. Los ejemplos que he dado proporcionan cierta evidencia de ello, pero daré más ejemplos en la proxima sección.

II

Algunas personas han sostenido —incluso antiguos alumnos míos— que mi teoría de la ciencia ha sido refutada por los hechos de la historia de la ciencia. Esto es un error: es un error

sobre los hechos de la ciencia y también es un error respecto a las afirmaciones de mi metodología.

Como traté de dejar claro en 1934 (*L. I. C.*, p. 48, y secciones 10 y 11) no considero la metodología como una disciplina empírica a contrastar, quizá, por medio de los hechos de la historia de la ciencia. Es, más bien, una disciplina filosófica, metafísica, y quizá, incluso, en parte, una propuesta normativa. Se basa en gran medida en el realismo metafísico y en la lógica de la situación: la situación de un científico sondeando la desconocida realidad que se esconde tras las apariencias y ansioso por aprender de los errores.

No obstante, siempre he pensado que mi teoría —de la refutación seguida por la aparición de un problema nuevo, seguida a su vez por una teoría nueva y quizá revolucionaria— era del mayor interés para el historiador de la ciencia, puesto que llevaba a una revisión del modo en que los historiadores debían juzgar la historia; especialmente ya que la mayoría de los historiadores creían, en aquellos días (1934), en una teoría inductivista de la ciencia[3] (ahora ya han renunciado a ella —incluso mis críticos—).

Que mi teoría, en la medida en que es exacta, tenga interés para los científicos y los historiadores no tiene nada de sorprendente; porque muchos de ellos —creo que la mayoría— comparten mi concepción realista del mundo y entienden los objetivos de la ciencia de la misma manera que yo: lograr explicaciones cada vez mejores.

Algunos ejemplos pueden resultar útiles.

Doy aquí una lista de casos interesantes en los que las refutaciones llevaron a reconstrucciones teóricas revolucionarias. Esta lista se remonta en gran parte a los años de 1930 y a mis tiempos en Nueva Zelanda, cuando di una serie de conferencias en la rama de Christchurch de la Royal Society of New Zealand, ilustrando mi teoría con ejemplos de la historia de la física. He escrito sobre algunos de esos casos en varios lugares; y no creo que esta lista contenga todos los casos de falsación a los que me he referido en mis diversos escritos. Al hacerlo, he confiado principalmente en mi memoria: no pretendo ser, yo mismo, un historiador de la ciencia. Y a causa de las presiones de otros trabajos urgentes nunca he tenido tiempo para repasar sistemáticamente la historia de la física en busca de otros ejemplos: no dudo que haya cientos. Pero creo que la lista que presento aquí —la lista de unos cuantos casos llamativos que sólo pueden entenderse como ejemplos de refutación— es suficientemente impresionante. (Incluso me sentiría inclinado a sugerir

[3] [Véase Joseph Agassi, *Towards a Historiography of Science*, 1963, *Ed.*]

que, históricamente, una ciencia se convierte en ciencia cuando ha aceptado una refutación empírica, pero realmente no propongo esto como hipótesis seria; y el caso de Copérnico *puede* ser un ejemplo de lo contrario: una gran teoría científica que no fue instigada por una refutación empírica.)

Una lista de ejemplos escogidos al azar

1) Parménides-Leucipo: Leucipo toma la *existencia del movimiento* como refutación parcial de la teoría de Parménides de que el mundo está lleno e inmóvil. Esto lleva a la teoría de «los atómos y el vacío». Es el fundamento de la teoría atómica[4].

2) Galileo refuta la teoría del movimiento, de Aristóteles: esto lleva a la creación de la teoría de la aceleración y, posteriormente, a la de las fuerzas newtonianas. Asimismo, Galileo toma las lunas de Júpiter y las fases de Venus como refutación de Ptolomeo[5] y, por tanto, como base empírica de la teoría rival de Copérnico.

3) Toricelli (y sus predecesores): la refutación «la naturaleza aborrece el vacío». Esto prepara el advenimiento de una teoría mecanicista del mundo.

4) La refutación de Kepler de la hipótesis del movimiento circular sostenida hasta entonces (incluso por Tycho y Galileo), lleva a las leyes de Kepler y, así, a la teoría de Newton[6].

---

[4] *Cf.* mi *El desarrollo del conocimiento científico: Conjeturas y Refutaciones* [*C. & R.*], Capítulo 2, secciones VI y VII. [Véase también el tomo III del *Post Scriptum*, «Epílogo metafísico». *Ed.*]

[5] En una carta muy interesante que me dirigió, Allan Franklin plantea algunas dudas sobre este problema, sugiriendo que, aunque Galileo arguyó en contra de la ley aristotélica del movimiento como si se refiriese sólo a velocidades constantes, el propio Aristóteles, así como sus comentaristas posteriores, reconocía que los cuerpos que caen se aceleran. Franklin encuentra que la cinemática del movimiento uniformemente acelerado fue elaborada en el siglo XIV en el Merton College, Oxford, y también en París (por Oresme). Se refiere también a la «teoría del ímpetu» de Buridan y al argumento de Domingo de Soto de que los cuerpos en caída libre ejemplifican la aceleración uniforme. Agradezco a Mr. Franklin estos comentarios. Véase también *El desarrollo del conocimiento científico*]*C. & R.*], Capítulo 3, sección 1.

[6] Los historiadores han afirmado con frecuencia que había un *prejuicio* en favor del movimiento circular que Kepler y otros tuvieron que vencer. Pero el movimiento circular no era simplemente un prejuicio; había, en efecto, una ley de conservación para la circularidad, no sólo para las rotaciones de los planetas, sino también para la rueda: la conservación del momento angular. Desde luego, las leyes de conservación no estaban claras en aquellos tiempos. Pero la ley de conservación de Galileo, su forma de la ley de la inercia, tomaba en consideración el movimientos circular; y todo esto se encontraba detrás de las ideas metafísicas y de los principios de explicación aceptados en aquel

5) La refutación de Lavoisier de la teoría del flogisto lleva a la química moderna.

6) La falsación de la teoría de la luz de Newton (el experimento de las dos ranuras de Young). Esto lleva a la teoría de la luz de Young-Fresnel. La velocidad de la luz en el agua en movimiento es otra refutación. Prepara para la relatividad especial.

7) El experimento de Oersted es interpretado por Faraday como una retufación de la teoría universal de las fuerzas centrales newtonianas y lleva así a la *teoría de campos* de Faraday y Maxwell.

8) Teoría atómica: la atomicidad del átomo es refutada por el electrón de Thomson. Esto lleva a la teoría electromagnética de la materia y, con el tiempo, a la aparición de la electrónica. Veánse los intentos de Einstein y de Weyl de encontrar una teoría monista («unificada») de la gravitación y la electromagnética[7].

9) El experimento de Michelson (1881-1887-1902, etc.) lleva a *Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Köpern* de Lorenz (1895: véase § 89). El libro de Lorenz tuvo una importancia crucial para Einstein, que aludió a él dos veces en § 9 de su artículo de 1905 sobre la relatividad. (El propio Einstein no consideraba muy importante el

---

momento. En este contexto, el movimiento circular era explicable, pero el movimiento elíptico se consideraba bastante irracional. Así, la *racionalidad* de la actitud metafísica (o programa metafísico de investigación) pre-kepleriana debería tenerse en cuenta: estaba la bien establecida continuación de la rotación de la rueda; si tenemos una rueda libremente suspendida, el momento angular está tan bien sustentado por la observación como lo están las fuerzas de inercia. A Kepler le costó gran esfuerzo superar esa concepción, pero no porque fuera un prejuicio; más bien porque formaba parte importante del trasfondo racional. Sintiendo, como sintió al principio, que el movimiento elíptico era irracional, Kepler necesitaba un nuevo tipo de explicación para él. Puede que fuese ahí donde entrase el sol: en la descripción de Kepler hay fuerzas que emanan del sol, mientras que en la teoría de Galileo el movimiento circular de los planetas no depende realmente del sol. Desde luego la teoría de Kepler es distinta de la nuestra: él hablaba, principalmente, no de la atracción ejercida por el sol, sino de que los rayos son empujados hacia fuera por el sol. Sólo con Newton empezó a verse claro que la atracción del sol influye a los planetas, de la misma manera que la atracción de la tierra influye a la luna. Galileo, sin embargo, seguía oponiéndose a la teoría de Kepler a causa de sus connotaciones astrológicas, es decir, sus connotaciones irracionales: la «influencia» de unos planetas sobre otros planetas. Kepler era, en efecto, un astrólogo; y la astrología mantiene que los cuerpos celestes ejercen fuerzas unos sobre otros. Así, uno puede entender tanto a Galileo como a Kepler. El marco metafísico de Galileo le imponía el movimiento circular y le impedía aceptar la influencia del sol y de la luna. Véase tambien *Conocimiento objetivo* [*Obj. Kn.*], Capítulo 4, sección 9.

[7] *Cf.* mi «The Rationality of Scientific Revolutions» en R. Harré, ed., *Problems of Scientific Revolution*, 1975, pp. 27-101; véase sección XII.

experimento de Michelson.) La teoría de la relatividad especial de Einstein es: *a*) un desarrollo del formalismo fundado por Lorenz, y *b*) una interpretación diferente —es decir, relativista— de ese formalismo. Hasta el momento no existe un experimento crucial para decidir entre las interpretaciones de Lorenz y de Einstein; pero si tenemos que adoptar la acción a distancia (no localidad: véase *Teoría cuántica y el cisma en física,* tomo III del *Post Scriptum*, Prefacio de 1982), entonces tendríamos que volver a Lorenz.

Por cierto, pasaron años antes de que los físicos empezaran a llegar a cierto acuerdo sobre la importancia de los experimentos de Michelson: yo no afirmo que normalmente se acepten las falsaciones en seguida (véase la sección precedente) —ni siquiera que se reconozcan inmediatamente como falsaciones *potenciales*—.

10) Los «descubrimientos fortuitos» de Roentgen y Becquerel refutaron ciertas expectativas (que se tenían inconscientemente), sobre todo las espectativas de Becquerel. Tuvieron, naturalmente, consecuencias revolucionarias.

11) La teoría de (parcial) éxito de Wilhelm Wien sobre la radiación de un cuerpo negro chocó con las teorías de Sir James Jeans y de Lord Rayleigh que también tuvieron mucho éxito (en parte, véase la sección precedente). La refutación de Lummer y Pringsheim de la fórmula de radiación de Rayleigh y Jeans, junto con la obra de Wien, lleva a la teoría cuántica de Planck (véase *L. I. C.*, p. 103). En esto, Planck refuta su propia teoría, la interpetación absolutista de la ley de la entropía, frente a una interpretación probabilística similar a la de Boltzmann.

12) Los experimentos de Philipp Lenard referentes al efecto fotoeléctrico estaban en conflicto, como insistía el propio Lenard, con lo que cabía esperar de la teoría de Maxwell. Llevaron a la teoría de Einstein de los cuantos de luz o fotones (que también estaban, naturalmente, en conflicto con Maxwell), y, así, mucho más tarde, al dualismo de onda y partícula.

13) La refutación de la teoría anti-atomista y fenomenalista de la materia, de Mach y Ostwald: el excelente artículo de Einstein sobre el movimiento browniano, en 1905, sugería que el movimiento browniano podría interpretarse como una refutación de esta teoría. Así pues, este artículo hizo mucho por establecer la realidad de moléculas y atómos.

14) La refutación de Rutherford del *modelo del átomo como vórtice*[8]. Esto lleva directamente a la teoría de Bohr de

[8] *Cf.* «The rationality of Scientific Revolutions», p. 90.

1913 sobre el átomo de hidrógeno y, por fin, a la mecánica cuántica.

15) La refutación de Rutherford (en 1919) de la teoría según la cual los elementos químicos no podían cambiarse artificialmente (aunque podrían desintegrarse espontáneamente).

16) La teoría de Bhor, Kramers y Slater (véase *L. I. C.*, p. 232): Esta teoría fue refutada por Compton y Simon. La refutación lleva casi inmediatamente a la mecánica cuántica de Heisenberg, Born y Jordan.

17) La interpretación de Schrödinger de su propia teoría (y de Broglie) es refutada por la interpretación estadística de ondas de materia (experimentos de Davisson y Germer y de George Thomson, por ejemplo). Esto lleva a la interpretación estadística de Born.

18) El descrubrimiento del positrón (1932) por Anderson refuta muchas cosas: refuta la teoría de dos partículas elementales —protones y electrones—; refuta la teoría de la conservación de las partículas; refuta la interpretación original de Dirac de su propia predicción de partículas positivas (él pensó que eran protones). Con ello se corrobora algún trabajo teórico de 1930-31. (Para más detalles, véase Norwood Russell Hanson, *The Concept of the Positron*, 1963; un libro excelente.)

19) La teoría eléctrica de la materia[9] elaborada por Einstein y Weyl, que Einstein sostuvo implícitamente —y, en cualquier caso, la buscó— hasta el fin de su vida (ya que él interpretaba la teoría unificada de campos como una teoría de *dos* campos, gravitatorio y electromágnetico), es refutada por el neutrón y por la teoría de Yukawa de las *fuerzas nucleares*: el Mesón de Yukawa. Esto da lugar a la teoría del núcleo.

20) La refutación de la conservación de la paridad (véase Allan Franklin, *Stud. Hist. Philos. Sci.* 10, 1979, p. 201).

III

Se entiende, naturalmente, que esas refutaciones no hicieron más que crear nuevas situaciones de los problemas que, a su vez, estimularon la imaginación y el pensamiento crítico. Las nuevas teorías que se desarrollaron no fueron, pues, resultados directos de las refutaciones: fueron los logros del pensamiento creador, de los hombres que pensaron.

Otra advertencia obvia es que, en algunos de esos casos, se tardó cierto tiempo antes de que la refutación se aceptase como

[9] *Cf.* la Introducción al tomo III del *Post Scriptum*. *Cf.* también «The Rationality of Scientific Revolutions», p. 90, primer párrafo nuevo.

tal: hubo con frecuencia acciones de retaguardia, a veces incluso acciones prolongadas, antes de que la refutación se aceptase en la práctica *como refutación* por todas las personas competentes, en vez de interpretarla de algún otro modo. Pero esto no fue siempre así: no lo fue, por ejemplo, en los casos 12) (los resultados de Lenard se aceptaron con bastante rapidez), 13) a 17) e incluso 18), aunque la teoría refutada tuvo, en este último caso, una larga supervivencia.

Hay, naturalmente, excepciones a este análisis en términos de refutación seguida de reconstrucción. La mayor excepción parece haber sido la de Copérnico, cuyo objetivo fue dar una explicación alternativa a los hechos explicados por Ptolomeo[10]. Para estar totalmente seguro de que ésta es realmente una excepción, habría que estudiar el caso con mayor detalle y, sobre todo, la *aceptación* de la teoría por los científicos, que pudo haberse retrasado hasta los nuevos descubrimientos empíricos de Galileo mencionados en 2), de los que puede afirmarse que son refutaciones de Ptolomeo.

Mi teoría de la ciencia no pretendía ser una teoría histórica ni ser una teoría apoyada en hechos históricos o empíricos de otro tipo, como he dicho antes. *Sin embargo, dudo que exista una teoría de la ciencia que pueda arrojar tanta luz sobre la historia de la ciencia como la teoría de que la refutación va seguida por una reconstrucción revolucionaria y, sin embargo, conservadora*[11].

---

[10] Lo que ocurrió en el caso de Copérnico fue que se reinterpretaron los hechos en términos de la vieja teoría de Aristarco, que había sido olvidada casi por completo, sin ninguna intención de hacer un experimento crucial. Copérnico quería decir que los mismos hechos pueden reinterpretase a la luz de la teoría de Aristarco. Sólo más tarde advirtieron otros que quizá los hechos pudieran interpretarse *mejor* a la luz de la teoría de Copérnico, y que había otras ventajas. Otra forma de expresar esto —aunque usando una terminología diferente de la que empleo Copérnico— es decir que propuso cambiar el trasfondo metafísico; y que la teoría de Ptolomeo era una teoría metafísica que sufría de dificultades graves: por ejemplo, el mecanismo de la envoltura es violado por los cometas que penetran en ella. Sólo cuando se reconoció que las dos teorías alternativas tenían consecuencias empíricas diferentes, se convirtió la cuestión en científica, o al menos, se convirtió en un programa de investigación vivo capaz de convertirse Estas contrastaciones fueron posibles de hecho con·Galileo: con las fases de Venus, las lunas de Júpiter y las diferencias de tamaño de Venus, Marte y Mercurio. (Los tamaños aparentes de los planetas tendrían que ser constantes si la tierra estuviese en el centro.) (Véase *Conocimiento objetivo* [*Obj. Kn.*, 1972], pp. 188-189.)

[11] Véase, por ejemplo, «The Rationality of Scientific Revolutions» en Rom Harré, ed., *Problemas of Scientific Revolution, op. cit.*, sección VIII, pp. 82 s.

IV

Este puede ser el lugar adecuado para mencionar, y refutar, la leyenda de que Thomas S. Kuhn, en su capacidad de historiador de la ciencia, es quien ha mostrado que mis concepciones sobre la ciencia (llamadas a veces, pero no por mí mismo, «Falsacion*ismo*») pueden refutarse con hechos, es decir, por medio de la historia de la ciencia.

No creo siquiera que Kuhn lo haya intentado. En todo caso, no lo ha logrado. Además, en la cuestión de la importancia de la falsación para la historia de la ciencia, las concepciones de Kuhn y las mías coinciden casi totalmente.

Esto no significa que no haya grandes diferencias entre las concepciones de Kuhn y las mías sobre la ciencia. Yo sostengo la antigua teoría de la verdad (casi explícita en Jenófanes, Demócrito y Platón y bastante explícita en Aristóteles), según la cual la verdad es la conformidad de lo que se afirma con los hechos. Las concepciones de Kuhn en esta cuestión fundamental me parecen estar afectadas por el relativismo; más específicamente, por una forma de subjetivismo y de elitismo como la propuesta, por ejemplo, por Polanyi. Kuhn me parece estar también afectado por el fideísmo de Polanyi: la teoría de que un científico *tiene* que tener fe en la teoría que propone (mientras que yo creo que los científicos —como Einntein en 1916 y Bohr en 1913— se hacen cargo a menudo de que están proponiendo conjeturas que tarde o temprano serán sustituidas por otras). Hay muchos otros puntos en los que diferimos, el más importante de los cuales es, quizá, mi énfasis en la crítica objetiva y racional: considero característico de la ciencia, antigua y moderna, el enfoque crítico de las teorías, desde el punto de vista de sin son verdaderas o falsas. Otro punto que me parece fundamental es que Kuhn no parece comprender la gran importancia de las muchas revoluciones puramente *científicas* que *no* están relacionadas con revoluciones *ideológicas*. En realidad, él casi parece identificar los dos tipos[12].

Pero con respecto a la falsabilidad o a la imposibilidad de obtener pruebas concluyentes de falsación y respecto a función de ambas en la historia de la ciencia y de las revoluciones científicas, no parece haber ninguna diferencia de importancia entre Kuhn y yo.

Kuhn, sin embargo, parece ver grandes diferencias entre nosotros en este punto, aunque también señala muchas simila-

[12] Véase «The Rationality of Scientific Revolutions», *op. cit.*, pp. 72-101, esp. 87-93.

ridades entre sus concepciones y las mías. Para explicar esas similaridades menciona que asistió a mis conferencias del ciclo «William James» en Harvard, en 1950. También era, podría añadir yo, uno de los miembros más activos y más críticos de mis seminarios (ochenta y dos horas de conferencias y ochenta y dos horas de seminarios, creo). Pero parece claro que no recuerda totalmente lo que ocurrió durante aquellas sesiones, y cuando escribió su primer libro, *The Copernican Revolution* (1957; edición en rústica, 1959), ya no guardaba, evidentemente, más que un recuerdo muy esquemático de mis concepciones y me consideraba un «falsacionesta ingenuo». Sin embargo, en ese libro, Kuhn aceptaba *prácticamente* mis auténticas concepciones sobre el carácter revolucionario de la evolución de la ciencia. Se desvía de mis concepciones sólo en que sostiene lo que he descrito antes como «fideísmo», porque afirma que «un científico *tiene que creer* [el subrayado es mío] en su sistema antes de confiar en él como guía para la investigación fructífera de lo *desconocido*»[13]. Pero Kuhn me sigue de cerca cuando continúa: «Pero el científico paga el precio de su compromiso... una sola observación incompatible con su teoría [puede demostrar][14] que ha estado empleando una teoría equivocada en todo momento. Entonces debe abandonar su esquema conceptual y sustituirlo»[15].

Esto es, obviamente, «falsacionismo»; de hecho, es algo parecido a un «estereotipo metodológico de la falsación» por citar la alusión que hace Kunh a mí en su libro posterior, *The Structure of Scientific Revolutions* (1962, p. 77). Pero en su libro anterior, sobre Copérnico, Kuhn continúa: «En esquema, ésta es la estructura lógica de una revolución científica. Un esquema conceptual... lleva finalmente a resultados que son incompatibles con las observaciones... Es un esquema útil, porque la incompatibilidad de teoría y observación es la fuente primaria de toda revolución en la ciencias».

Este «esquema útil» no es sólo falsacionista; es un estereotipo mucho más simplista del falsacionismo que nada de lo que yo mismo haya dicho jamás en mis escritos, en mis conferencias o mis seminarios; en realidad, siempre he estado en total

---

[13] T. S. Kuhn, *The Copernican Revolution*, pp. 75 ss. de la edición en rústica, 1959. He subrayado las palabras «debe creer», porque el fideísmo es el único punto en este pasaje en el que Kuhn se desvía de mí: yo hubiera dicho que el científico *puede* creer; o, alternativamente, que puede aceptar «su sistema» sólo provisionalmente (como sabemos que hizo Einstein, por ejemplo; y Niels Bohr, por lo menos antes de 1926).

[14] *Loc. cit.*, Kuhn escribe «demuestra».

[15] No me gusta el término «su esquema conceptual»; yo hubiera dicho «su teoría».

acuerdo con la siguiente observación, más crítica, que añade Kuhn:

> Pero, históricamente, el proceso de la revolución no es nunca [yo diría: «casi nunca»], y no podría ser de ningún modo [Rutherford, ¿ve usted esto?] tan sencillo como indica el esquema lógico. Como ya hemos empezado a descubrir, la observación no es nunca *absolutamente* incompatible con una [teoría].

Naturalmente, yo hacía hincapié en este punto ya en 1934 (yo había señalado siempre que «la observación está impregnada de teoría», de la misma manera que señalaba que es imposible presentar una «contra-prueba» incuestionable de una teoría empíricamente científica. Véase la sección I más arriba). Por ello, me quedé perplejo cuando leí en el segundo libro de Kuhn, *The Structure of Scientific Revolutions* (p. 77):

> Ningún proceso de los revelados hasta el momento por el estudio histórico del desarrollo científico *se asemeja en absoluto* al estereotipo metodológico de la falsación por comparación directa con la naturaleza. [El subrayado es mío.]

¿Qué quería decir Kuhn con esto? ¿Que el proceso histórico no se asemeja en absoluto a un proceso de falsación, o que no se asemeja a ese «estereotipo» que él caracteriza como la «comparación directa con la naturaleza» que él llama en otro lugar «falsacionismo ingenuo» y que yo, por lo menos, siempre he rechazado?

Ahora resulta que es en mi propia teoría en la que Kuhn está pensando cuando habla de un «estereotipo» de la falsación. Porque en otro lugar (en P. A. Schilpp, ed., *The Philosophy of Karl Popper*, p. 808), escribe:

> Sir Karl no es, desde luego, un falsacionista ingenuo. Sabe todo lo que acabamos de decir y lo ha resaltado desde el comienzo de su profesión... *Aunque no es un falsacionista ingenuo, sugiero que puede considerarse a sir Karl legítimamente como tal.* [El subrayado es mío.]

Este es un pasaje realmente sombroso. Es lo mismo que decir: «Aunque Popper no es un asesino, yo sugiero que puede considerársele legítimamente como tal».

No hay auténticos argumentos en el artículo de Kuhn que conduzcan a este asombroso veredicto; excepto que él cree que mis propios argumentos contra lo que él llama «falsacionismo ingenuo» amenazan la «integridad de (mi) posición básica». Pero lo que él cree, erróneamente, mi «posición básica» es solamente *la leyenda, o paradigma*, según el cual Popper es un falsacionista ingenuo. Este argumento (si puede llamársele así) es circular.

Creo que Kuhn cree de verdad lo que escribe. Pero ¿cómo puede creerlo? He probado varias teorías explicativas. Sólo una

de ellas me parece plausible. Es que Kuhn, al principio de su profesión, se formó una teoría de mis concepciones que se convirtió en su paradigma de Popper: Popper es el hombre que sustituyó el verificacionismo por el falsacionismo («ingenuo»). Kuhn formó este paradigma (siguiendo sus propias indicaciones) antes de haber leído ninguno de mis escritos. Cuando por fin leyó *La lógica de la investigación científica*, lo leyó a la luz de su paradigma. Muchos pasajes de este libro (uno en la página inmediatamente siguiente a mi introducción de la idea de falsación) mostraban que yo no me ajustaba a su paradigma. Pero, como hemos aprendido de Kuhn, no es fácil renunciar a los paradigmas[16].

La cuestión ahora es ésta: ¿Soy yo realmente el hombre que tenía el falsacionismo ingenuo como pieza clave de su pensamiento? ¿Es verdadero el paradigma de Kuhn? ¿Puedo ser «legítimamente considerado como un falsacionista ingenuo», a pesar de que Kuhn admite después de leer *La lógica de la investigación científica* que, ya en 1934, yo *no* lo era?

Resulta que la verdadera pieza clave de mi pensamiento sobre el conocimiento humano el es falibilismo y el enfoque crítico; y que veo, y vi incluso antes de 1934 (véase mi *Die beiden Grundprobleme der Erkenntnistheorie*) que el conocimiento humano es un caso muy especial de conocimiento animal. Mi idea central en el terreno del conocimiento animal (incluido el conocimiento humano) es que se basa en conocimiento heredado. Tiene el carácter de las expectaciones inconscientes. Siempre se desarrolla como consecuencia de modificaciones del conocimiento anterior. La modificación es (o es como) una mutación: viene de dentro, tiene el carácter de un globo de prueba, es intuitiva o atrevidamente imaginativa. Tiene, pues, un carácter conjetural: la expectación puede resultar frustrada, el globo o la pompa pueden pincharse: toda la información que se recibe de fuera es eliminatoria, selectiva.

Lo especial del conocimiento *humano* es que puede formu-

---

[16] La contribución de Kuhn a P. A. Schilpp, ed., *The Philosophy of Karl Popper* (tomo II, pp. 798-819) es una crítica agradablemente escrita del legendario falsacionista ingenuo K. R. P. Está tan convencido de conocer mis opiniones y sus puntos débiles que, con mis libros en las manos, me habla de «locuciones» tales como «falsación» o «refutación» que son «antónimos de prueba». Pero en el Indice de Materias de mi *L. Sc. D.* podría haber encontrado: «Contra-prueba, no puede ofrecerse una contra-prueba concluyente (de una teoría), 42, 50, 81-87». [a][Esta referencia es a la versión inglesa. No se encuentra en la española][a] Esta y otras muchas observaciones en su contribución muestran lo que le ocurre al lector de un libro tiene un «paradigma» de lo que debe encontrarse en él y de lo que no. En general, encuentro que muchos historiadores de la ciencia son muy malos lectores (es decir, están llenos de prejuicios).

larse en un lenguaje, por medio de proposiciones. Ello hace posible que el conocimiento se haga consciente y sea criticable objetivamente por medio de argumentos y de contrastaciones. De este modo llegamos a la ciencia. Las contrastaciones son intentos de refutación. Todo conocimiento sigue siendo falible, conjetural. No hay justificación ni, naturalmente, justificación definitiva de una refutación. No obstante, aprendemos por las refutaciones, es decir, por la eliminación de errores, por un proceso de realimentación[17]. En esta descripción no hay lugar en absoluto para la «falsación ingenua».

V

Otra objeción a mi teoría del conocimiento está mejor fundada, aunque su impacto en mi teoría es insignificante. Es el admitido fracaso de una definición (de la similitud con la verdad o aproximación a la verdad) que propuse en 1963[18].

Voy a explicar primero con dos ejemplos, 1) y 2), el único tipo de uso de la idea de similitud con la verdad que puede aparecer en mi teoría del conocimiento (o en la de cualquier otro).

1) El enunciado de que la tierra está en reposo y que los cielos estrellados rotan a su alrededor está más lejos de la verdad que el de que la tierra rota alrededor de su propio eje; que es el sol el que está en reposo; y que la tierra y los otros planetas giran en órbitas circulares alrededor del sol (como propusieron Copérnico y Galileo). El enunciado, debido a Kepler, de que los planetas no se mueven en círculos, sino en elipses (no muy alargadas) con el sol en su foco común (y con el sol en reposo o girando sobre su eje) es una mayor aproximación a la verdad. El enunciado, debido a Newton, de que existe un espacio en reposo, pero que, aparte de la rotación, su posición no puede hallarse por medio de la observación de las estrellas o de efectos mecánicos, en un paso más en la dirección de la verdad.

2) Las ideas de Gregor Mendel sobre la herencia estaban más próximas a la verdad, al parecer, que las concepciones de Charles Darwin. Los experimentos posteriores de reproducción con moscas de la fruta (mosca mediterránea) llevaron a mayores progresos en la similitud con la verdad de la teoría de la herencia. La idea del acervo genético de una población (una especie) supuso un paso más. Pero los pasos más importantes,

[17] No creo que Norbert Wiener, en su *Cybernetics*, se refiriese al darwinismo o a la eliminación de errores (prueba y error).

[18] Véase *El desarrollo del conocimiento científico* [*C. & R.*], Capítulo 10.

con mucho, fueron los que culminaron con el descubrimiento del código genético.

Estos ejemplos, 1) y 2), muestran, creo, que no es necesaria una definición formal de similitud con la verdad para poder hablar inteligentemente de ella (véase también, más adelante, el *Addendum* al capítulo IV, Parte I).

Entonces, ¿por qué intenté dar una definición formal?

He argüido a menudo que las definiciones son innecesarias. Nunca son realmente necesarias y rara vez útiles, excepto en el siguiente tipo de situación: introduciendo una definición podemos mostrar que no sólo se necesitan menos suposiciones básicas para una buena teoría, sino que, además, nuestra teoría puede explicar más que sin la definición. En otras palabras, una nueva definición tiene interés sólo si la teoría resulta fortalecida por ella. Creí que podía hacer esto con mi teoría de los objetivos de la ciencia: la teoría de que la ciencia busca la verdad y la resolución de problemas de explicación, es decir, que busca teorías de mayor capacidad explicativa, mayor contenido y mayor contrastabilidad. La esperanza de fortalecer aún más esta teoría de los objetivos de la ciencia por medio de la definición de similitud con la verdad en términos de verdad y de contenido fue, desgraciadamente, vana. Pero la concepción, muy extendida, de que el abandono de esta definición debilita mi teoría, carece por completo de fundamento. Puedo añadir que acepté la crítica hecha a mi definición a los pocos minutos de su presentación, preguntándome cómo no habría visto el error yo mismo antes, pero nadie ha mostrado nunca que mi teoría del conocimiento, que desarrollé ya en 1933 y que lleva creciendo vigorosamente desde entonces y que es muy utilizada por los científicos en activo, haya resultado debilitada en lo más mínimo por esta infortunada definición errada o que la idea de similitud con la verdad (que no es una parte esencial de mi teoría) no pueda seguir usándose dentro de mi teoría como concepto sin definir.

La aserción de que mi *autoridad* ha resultado dañada por este incidente es obviamente cierta, pero nunca he prentendido ni deseado tener ninguna autoridad. La aserción de que *mi teoría* ha resultado dañada se ha lanzado sin siquiera intentar dar una razón, y me parece que es una afirmación incompetente.

VI

También se le ha encontrado a mi teoría la objeción de que no puede responer a la paradoja de Nelson Goodman[19].

[19] [Véase W. W. Bartley III, «Eine Lösung des Goodmen-Paradosons» en

Que no es así se verá en las siguientes consideraciones que muestran, por medio de un simple cálculo, que el enunciado de evidencia *e* «todas las esmeraldas observadas antes del 1 de enero del año 2000 son verdes» no hace a la hipótesis h_1, «todas las esmeraldas son verdes, al menos hasta febrero del año 2000» más probable que las hipótesis «todas las esmeraldas son azules, para siempre jamás, con la excepción de las que fueron observadas antes del año 2000, que son verdes». Esto no es una paradoja que hay que formular y disolver por medio de investigaciones lingüísticas, sino que es un teorema demostrable del cálculo de probabilidades. El teorema puede formularse como sigue:

El cálculo de probabilidades es incompatible con la conjetura de que la probabilidad es ampliativa (y, por tanto, inductiva).

La idea de que la probabilidad es ampliativa está muy extendida. Es la idea de que la evidencia *e* —por ejemplo, que todos los cisnes en Austria son blancos— aumentará de algún modo la probabilidad de un enunciado que va más allá de *e*, tal como h_2 «todos (o la mayoría de) los cisnes en las regiones limítrofes de Austria son blancos». Dicho de otro modo, la idea es que la evidencia hace al menos un poco más probables cosas que están *más allá* de lo que realmente afirma. (Esta concepción fue defendida con fuerza por Carnap, por ejemplo.)

Fue especialmente el teorema que damos a continuación el que sugirió la concepción según la cual la probabilidad es ampliativa (*h* = hipótesis; *e* = evidencia empírica; *b* = información básica):

Sea $p\,(h,b) \neq 0$. Además, sea *e* evidencia favorable (es decir, *e* se sigue de *h* en presencia de *b*, de modo que $p\,(e,b) \neq 1$ y $p\,(e,\ hb) = 1$). Entonces $p\,(h,eb) > p\,(h,b)$. Esto es, la evidencia favorable *e* hace *h* más probable, aunque *h* dice más que *e*. Y esto es válido para cada nueva e_1, e_2, ..., que satisfaga las mismas condiciones.

Parece, por tanto, que una evidencia favorable en aumento sigue sustentando a *h*; y de ese modo, parece que el apoyo es ampliativo.

Pero eso es una ilusión, como puede mostrarse de la manera siguiente:

Sean h_1 y h_2 dos hipótesis cualesquiera apoyadas por *e* en presencia de *b*, de modo que

$$p\,(e,b) \neq 1 \text{ y } p\,(e,\ h_1b) = p\,(h_2b) = 1$$

G. Radnitzky und Gunnar Andersson, *Voraussetzungen und Grenzen der Wissenschaft* (Tübingen, 1982), pp. 347-358; y en «Rationality, Criticism and Logic», *Philosophia* 11, febrero 1982, esp. pp. 169-173, y las referencias que allí se dan. *Ed.*]

Sea $R_{1,2}$ *(anterior)* = $p\,(h_1b) \,/\, p\,(h_2b)$ el cociente de las probabilidades de h_1 y h_2 con anterioridad a la evidencia e, y sea

$$R_{1,2}\ (posterior) = p\,(h_1,\ eb) \,/\, p\,(h_2,eb)$$

el cociente de las dos probabilidades con posterioridad a la evidencia e.

Entonces tenemos, para una h_1, h_2, y e cualesquiera que satisfagan dichas condiciones:

$$R_{1,2}\ (anterior) = R_{1,2}\ (posterior)$$

Esto se sigue casi inmediatamente de

$$p\,(a,bc) = p\,(a,bc) \,/\, p\,(b,c)$$

es decir, del teorema de Bayes.

¿Qué significa

$$R_{1,2}\ (posterior) = R_{1,2}\ (anterior)?$$

Dice que la evidencia no cambia el cociente de las probabilidades anteriores, tanto si las hemos calculado como si las hemos supuesto libremente, siempre que las dos hipótesis puedan explicar, ambas, la evidencia e. Pero eso significa que si suponemos que:

h_1 = todos los cisnes en *alguna* región más grande que Austria son blancos;

h_2 = todos los cisnes del mundo son no blancos, excepto los de Austria, que son blancos;

e = todos los cisnes en Austria son blancos,

entonces, suponiendo cualquier probabilidad anterior que se desee para h_1 y h_2: su cociente $R_{1,2}$ *(anterior)* permanece inalterado por la evidencia. Así pues, no hay extensión, no hay apoyo ampliativo: no hay probabilidad ampliativa, ni para los cisnes ni para las esmeraldas. Y esto no es absurdo, sino tautológico (y no lo afectan las traducciones).

REALISMO Y EL OBJETIVO DE LA CIENCIA

Esto hay de cierto: que nada hay tan adecuado para formar una mente capaz de un gran desarrollo como vivir y participar en grandes revoluciones científicas. Yo aconsejaría, por tanto, a todos aquellos a quienes la época en que viven no les ha deparado naturalmente este beneficio, que se lo proporcionen artificialmente, con la lectura de los escritos de las épocas en que las ciencias han sufrido grandes cambios. Leer los escritos de los sistemas más opuestos y extraer su verdad oculta, responder a las preguntas suscitadas por esos sistemas opuestos, transferir las teorías principales de un sistema a otro, es un ejercicio que nunca le será suficientemente recomendado a un estudiante. Su esfuerzo se vería ciertamente recompensado, ya que llegaría a ser tan independiente de la estrechez de las opiniones de su propio tiempo como es posible.

HANS CHRISTIAN OERSTED[1]

Los matemáticos pueden hacerse la ilusión de que poseen ideas nuevas que el simple lenguaje humano es todavía incapaz de expresar. Que hagan el esfuerzo de expresar esas ideas con palabras apropiadas, sin la ayuda de símbolos y, si lo logran, no sólo quedaremos nosotros, los profanos, en deuda permanente con ellos, sino que, nos atrevemos a decir, ellos mismos saldrán muy ilustrados del proceso e incluso dudarán si las ideas expresadas en símbolos habían encontrado realmente alguna vez el camino desde las ecuaciones a sus mentes.

JAMES CLARK MAXWELL[2]

---

[1] Hans Christian Oersted, «Observations on the History of Chemistry: A Lecture, 1805-1807» en *The Soul in Nature, with Supplementary Contributions* (Londres: Henry G. Bohn, 1852), p. 322.

[2] W. D. Niven, ed., *Scientific Papers of James Clark Maxwell*, Vol. II (Cambridge, 1890), p. 328; reimpreso de *Nature* 7, 27 marzo 1873, p. 400: Reseña de la obra de Thomson y Taite, *Elements of Natural Philosophy*.

PREFACIO DE 1956

SOBRE LA INEXISTENCIA DEL METODO CIENTIFICO

Mas, en realidad, nada sabemos por haberlo visto; porque la verdad está oculta en lo más profundo.

DEMÓCRITO[1]

Por regla general, empiezo mis clases sobre el Método Científico diciendo a mis alumnos que el método científico no existe. Añado que debo saberlo, porque he sido, al menos durante cierto tiempo, el solo y único profesor de esta disciplina inexistente, en toda la Commonwealth Británica.

Hay varios sentidos en los que no existe mi disciplina, y voy a mencionar algunos.

En primer lugar, mi disciplina no existe porque, en general, las disciplinas no existen. No hay disciplinas; no hay ramas del saber o, más bien, de la investigación: sólo hay problemas, y el impulso de resolverlos. Una ciencia tal como la botánica o la química (o, digamos, la fisicoquímica o la electroquímica) es, sostengo, una mera unidad administrativa. Los administradores de las universidades tienen un trabajo difícil, de todos modos, y les resulta muy conveniente trabajar con el supuesto de que existen ciertas disciplinas con sus nombres, con cátedras vinculadas a ellas, a ocupar por expertos en esas disciplinas. Se ha dicho que las asignaturas son también útiles para los estudiantes. No estoy de acuerdo: incluso los estudiantes serios se dejan engañar por el mito de la asignatura. Y no estoy dispuesto a decir que una cosa que engaña a una persona es una cosa útil para ella.

Eso en cuanto a la inexistencia de disciplinas en general. Pero el Método Científico ocupa una posición algo peculiar al ser incluso menos existente que otras disciplinas inexistentes.

Lo que quiero decir es esto: los fundadores de la disciplina, Platón, Aristóteles, Bacon y Descartes, así como la mayoría de

[1] Véase Hermann Diels, *Die Fragmente der Vorsokratiker*, ed. Walther Kranz, 6 Auflage, 1951, vol. II, p. 166; 68 B 117.

sus sucesores, por ejemplo John Stuart Mill, creyeron que existía un método para encontrar la verdad científica. En un período posterior y ligeramente más escéptico hubo metodólogos que creían que existía un método, si no para *encontrar* una teoría verdadera, al menos para averiguar si una hipótesis *dada* era o no verdadera: o (aún más escépticos) si una hipótesis dada era al menos «probable» en cierto grado averiguable.

Yo afirmo que no existe método científico en ninguno de esos tres sentidos. Para expresarlo de forma más directa:

1) No existe método para descubrir una teoría científica.

2) No existe método para cerciorarse de la verdad de una hipótesis científica, es decir, no existe método de verificación.

3) No existe método para averiguar si una hipótesis es «probable» o probablemente verdadera.

Una vez que he explicado así a mis alumnos que no hay nada que pueda llamarse método científico me apresuro a comenzar mi disertación y trabajamos mucho. Porque un año apenas es suficiente para arañar la superficie de una disciplina, incluso inexistente.

¿Qué enseño a mis alumnos? Y ¿cómo puedo enseñarles?

Yo soy un racionalista. Por racionalista entiendo un hombre que quiere comprender el mundo y aprender mediante la discusión con otros. (Noten que no digo que un racionalista sostiene la teoría errónea de que los hombres son total o principalmente racionales.) Por «discutir con otros» entiendo, más concretamente, criticarlos, suscitar sus críticas y tratar de aprender de ellas. El arte de la discusión es una forma peculiar del arte de la lucha, con palabras en vez de espadas, e inspirado por el interés de acercarse a la verdad sobre el mundo.

No creo en la actual teoría de que para que una discusión sea fructífera, los que discuten tengan que tener mucho en común[2]. Por el contrario, creo que cuanto más distinta sea su formación, más fructífera será la discusión. Ni siquiera hace falta empezar por tener una lengua común: si no hubiera existido la Torre de Babel hubiéramos tenido que construirla. La diversidad hace más fructífera la discusión crítica. Las únicas cosas que deben tener en común los participantes en una discusión son el deseo de saber y la disposición a aprender del otro, mediante la crítica severa de sus concepciones —en la versión más fuerte que sea posible dar a sus concepciones— y escuchando lo que él tiene que decir como réplica.

---

[2] [Véase «The Myth of the Framework», en Eugene Freeman, ed., *The Abdication of Philosophy: Philosophy and the Public Good. Essays in Honor of Paul Arthur Schilpp*, 1976, pp. 23-48; y «Addendum: Hechos, normas y verdad: Otra crítica del relativismo» en *La Sociedad abierta y sus enemigos* [*O. S.*, 1962]. *Ed.*]

Creo que *el llamado método de la ciencia consiste en este tipo de crítica.* Las teorías científicas se distinguen de los mitos simplemente en que pueden criticarse y en que están abiertas a modificación a la luz de las críticas. No pueden ni verificarse ni probabilificarse.

Mi actitud crítica —o, si ustedes prefieren, herética— influye, naturalmente, en mi actitud respecto a mis colegas filósofos.

Todos conocerán la historia del soldado que descubrió que nadie en su batallón (menos él, claro) llevaba el paso. Yo mismo me encuentro constantemente en esta divertida posición. Y tengo mucha suerte porque, por lo general, algunos de los otros miembros del batallón están bastante bien dispuestos a llevar el paso. Esto aumenta la confusión; y como no soy un admirador de la disciplina en filosofía, estoy satisfecho mientras haya bastantes miembros del batallón que hayan perdido el paso suficientemente los unos respecto a los otros.

Algunas de las cosas que me hacen perder el paso y que me gusta criticar son:

1) *Las modas:* no creo en las modas, corrientes de opinión, tendencias ni escuelas, ni en ciencia, ni en filosofía. De hecho, creo que la historia de la humanidad podría describirse muy bien como una historia de erupciones de enfermedades filosóficas y religiosas de buen tono. Estas modas sólo tienen una función seria: la de provocar críticas. No obstante creo en la tradición racionalista de una república del saber y en la necesidad urgente de salvar esta tradición.

2) *La imitación de la ciencia física:* Me disgusta el intento que se hace en campos ajenos a las ciencias físicas, de imitarlas usando sus supuestos «métodos»: medición e «inducción a partir de la observación». La doctrina de que hay tanta ciencia en una disciplina como matemáticas contiene, o tanta ciencia como medición o «precisión» hay en ella, descansa en un total malentendido. Por el contrario, esta máxima es válida para todas las ciencias: Nunca busquéis más precisión de la que requiere el problema que tenéis entre manos.

Así, no tengo fe en la precisión: creo que la simplicidad y la claridad son valores en sí mismas, pero no que la precisión o la exactitud sea un valor en sí misma. La claridad y la precisión son objetivos diferentes y, a veces, incluso incompatibles. No creo en lo que a menudo se ha llamado «terminología exacta»: no creo en las definiciones[3], y no creo que las definiciones aumenten la exactitud; y me disgusta en particular la termino-

[3] [Véase *La sociedad abierta y sus enemigos* [*O. S.*, 1945], Capítulo 11; y *Búsqueda sin término* [*U. Q.*], sección 7. *Ed.*]

logía pretenciosa y la pseudo-exactitud que se ocupa de ella. Lo que puede decirse, puede y debe decirse siempre de forma cada vez más clara y más sencilla.

3) *La autoridad del especialista:* No creo en la especialización ni en los expertos. Con tener excesivo respeto al especialista estamos destruyendo la república del saber, la tradición racionalista y la propia ciencia.

Para concluir, creo que no hay más que un camino hacia la ciencia o hacia la filosofía, en este sentido: encontrar un problema, ver su belleza y enamorarse de él; casarse con él y vivir feliz con él, hasta que la muerte os separe, a menos que uno encuentre otro problema más fascinante aún, o a menos, naturalmente, que encuentre una solución. Pero incluso si uno encuentra solución, puede descubrir entonces, para su delicia, la existencia de toda una familia de encantadores, aunque quizá difíciles, problemas hijos por cuyo bienestar puede trabajar, con un objetivo, hasta el fin de sus días[4].

[4] [Este «Prefacio de 1956» se leyó en una reunión de The Fellows of the Center for Advanced Study in the Behavioral Sciences en Stanford, California, en noviembre de 1956. *Ed.*]

PARTE I
EL ENFOQUE CRITICO

CAPITULO PRIMERO

LA INDUCCION

Y, en cuanto a la verdad cierta,
no la ha conocido el hombre
ni la conocerá; ni de los dioses
ni aun de las cosas de que hablo.
Y, si por azar, la verdad perfecta
revelase, él mismo no lo sabría:
Pues todo es una red de conjeturas.

JENÓFANES

En este capítulo introductorio trataré el problema de la inducción de modo más completo y en un contexto más amplio que en mi *Lógica de la investigación científica* (*L. I. C.* en abreviatura), obra a la que ésta sirve de corolario: aquí discutiré todas las ramificaciones más interesantes que conozco del problema.

En la sección 2 trato de dar un esquema de mi teoría del conocimiento para reformular el problema de la inducción y para re-enunciar su solución. La discusión de las concepciones —en gran medida metafísicas— que tienden a impedir que se acepte esta solución, se continúa hasta la sección 16, titulada «Dificultades del realismo metafísico. Por un realista metafísico».

El capítulo II (secciones 17 a 26) se ocupa del *problema de la demarcación*, la demarcación entre la ciencia y la metafísica. (No trato de deslindar el sentido y el sin sentido.) Allí trato de mostrar que el problema de la demarcación y su solución por medio de un *criterio de demarcación según la contrastabilidad* tienen una significación que va mucho más allá de los límites de la filosofía.

El capítulo III (secciones 27 a 32) se ocupa del problema de la corroboración y de introducir el término técnico de «*grado de corroboración*». Trato mostrar que este problema tiene cierto interés en vista de la situación existente en la filosofía de la ciencia con respecto a los problemas; porque su solución corrige la creencia, muy divulgada pero *errónea*, de que la

inducción científica puede ayudarnos a evaluar la probabilidad de una hipótesis.

1. UN FILOSOFO DESCONCERTADO EN EL EXTRANJERO

Poco después de llegar a Londres por primera vez, en el otoño de 1935, aproximadamente un año después de la publicación de *Logik der Forschung (L. d. F.* en abreviatura), me llevaron a una reunión de la Sociedad Aristotélica. Bertrand Russell, a quien yo admiraba desde hacía tiempo como el filósofo más grande desde Kant, leyó un artículo sobre «Los límites del empirismo»[1]. Creyendo que nuestro conocimiento empírico se obtenía por inducción y profundamente impresionado por la fuerza de la crítica de Hume, Russell sugería que tendríamos que suponer algún *principio de inducción* que, a su vez, no podría basarse en la inducción; un principio cuya adopción marcase los límites del empirismo. La posición de Russell era la misma que yo había atribuido, con razón o sin ella, a Kant, en la primera sección de mi *L. d. F.*

Me invitaron a participar en el debate y dije que no creía en absoluto en la inducción, aunque creía en el empirismo, un empirismo que no imponía sobre sí mismo aquellos límites kantianos que Russel estaba dispuesto a aceptar. Esta afirmación (que formulé tan sarcásticamente como pude con el poco inglés de que disponía) fue tomada como una broma por el público, que se rió y aplaudió con benevolencia. Sugerí entonces que todo el problema venía de la creencia errónea de que el *conocimiento científico* era un tipo de conocimiento especialmente estricto o seguro o augusto. Esta declaración recibió la misma acogida que la primera. Concluí con un intento de explicar que en el sentido usual de saber, cuando yo *sé* que llueve, tiene que ser *verdad* que está lloviendo; porque si no es verdad, entonces, simplemente, no puedo saber que está lloviendo, por muy sinceramente que crea que lo sé. En este sentido de la palabra, «saber» siempre significa «conocimiento verdadero y seguro»; y «saber» significa, además, estar en posesión de *suficiente razón* para mantener que nuestro conocimiento es verdadero y seguro. Pero, dije, no existe un conocimiento científico en este sentido. Si, a pesar de ello, decidimos dar a los resultados de nuestros esfuerzos científicos el nombre habitual de «conocimiento científico», deberíamos tener claro, entonces, que el

[1] Publicado en *Proccedings of the Arist. Soc.* 36, 1936, pp. 131-150. Mis observaciones aludían especialmente a pp. 146 ss. [Véase Popper, *Búsqueda sin término* (*U. Q.*, 1976), sección 22. *Ed.*]

conocimiento científico no es una especie de conocimiento; y menos aún una especie distinguida por un elevado grado de solidez o de certeza. Por el contrario, medido con los elevados criterios de la crítica científica, el «conocimiento científico» continúa siendo siempre pura conjetura, aunque conjetura controlada por medio de la crítica y el experimento. (Ni siquiera podría alcanzar un grado positivo de «probabilidad» si se utilizase este término en el sentido del cálculo de probabilidades en la forma que le dieron, por ejemplo, Keynes o Jeffreys.) Terminé diciendo que, simplemente con admitir que el conocimiento científico consiste en conjeturas o hipótesis, podemos resolver el problema de la inducción sin tener que suponer un principio de inducción ni ningún límite del empirismo.

Mi pequeño discurso fue bien recibido, lo que fue grato, pero por razones equivocadas, lo que era desconcertante. Porque lo habían tomado como un ataque a la ciencia o quizá, incluso, como expresión de una actitud de superioridad frente a ella.

Hay que admitir que, por implicación, yo había atacado a la Ciencia con mayúscula y a aquellos de sus devotos dispuestos a tomar sus aserciones como la verdad del evangelio. Pero yo sabía, naturalmente, en Russell, con su profunda y crítica comprensión de la ciencia y con su amor a la verdad, no era uno de esos devotos. Así, el rechazo de la Ciencia con mayúscula, aunque implícito en lo que había dicho yo, no había sido el principal punto de mi argumento. Más bien lo que yo había querido transmitir al público era esto: si suponemos que lo que llamamos «conocimiento científico» consiste sólo en suposiciones o conjeturas, entonces este supuesto es suficiente para resolver el problema de la inducción —al que Kant llamó «el problema de Hume»— sin sacrificar al empirismo; es decir, sin adoptar un principio de inducción y sin atribuirle validez *a priori*. Porque *las suposiciones no se «inducen a partir de las observaciones»* (aunque *pueden*, desde luego, sernos sugeridas por las observaciones). Este hecho nos permite aceptar sin reservas (y sin los límites de Russell al empirismo) la crítica lógica de Hume a la inducción y renunciar a la búsqueda de una lógica inductiva, a la búsqueda de certidumbre, e incluso de probabilidad, mientras seguimos nuestra búsqueda científica de la verdad.

Esto, mi argumento principal, se perdió. Y me di cuenta de que difícilmente podía haber sido de otro modo. Porque si la gente piensa siguiendo unas pautas inductivas —¿y quién no lo hace?— entonces, una observación como «no creo en la inducción» no puede interpretarse más que en el sentido de «no creo en la ciencia». Tampoco creo que hubiese trasmitido mejor lo

que quería si hubiera empezado con las palabras «creo en la grandeza de la ciencia, pero no creo que los métodos y los procedimientos de la ciencia sean inductivos en ningún sentido». Si hubiera dicho esto, la gente hubiese oído «Ciencia» en vez de «ciencia» y habrían llegado a la conclusión de que yo quería defender alguna doctrina de la intuición o intuicionismo o, quizá, alguna forma de autoritarismo científico.

Puesto que acabo de utilizar dos «ismos», tal vez deba mencionar que me han reprochado con frecuencia, incluso algunos de mis amigos filosóficos más favorables, la mala costumbre de discutir la filosofía en términos «ismos»; y estoy dispuesto a admitir que sería más directo si uno pudiera explicar sus principios, enunciar sus argumentos y terminar, sin necesidad de referirse a «ismos». Pero mi desconcertante experiencia ilustra por qué este método no siempre funciona. Nunca nos dirigimos a mentes completamente abiertas. Por muy abierto que pueda ser un público, no puede dejar de albergar, aunque sea inconscientemente, teorías relacionadas, concepciones y expectaciones sobre el mundo e incluso sobre las formas por las que aprendemos a conocerlo. Tienen posiciones adoptadas; normalmente, posiciones típicas: «ismos».

La mayoría de nosotros, sobre todo la mayoría de los filósofos, tenemos gran cantidad de teorías de modo consciente y tras un examen crítico; y podemos estar preparados a defenderlas con argumentos y a renunciar a ellas cuando se presenten buenos argumentos en contra de ellas. Pero podemos tener también teorías que damos por sentadas de un modo más o menos inconsciente y, por tanto, acríticamente; y esas teorías que se tienen acríticamente, a veces contienen la razón más fuerte para continuar sosteniendo las otras teorías conscientemente. Que esto es así se sabe desde hace mucho tiempo: Bacon describió esos supuestos inconscientes como ídolos y *prejuicios.* En los diálogos de Platón, Sócrates hace con frecuencia que sus interlocutores se den cuenta de que ciertas posiciones adoptadas por ellos implican que tienen teorías o concepciones de las cuales no son plenamente conscientes y que, en ocasiones, incluso, están en conflicto entre sí. Antes de él, Parménides habla de las opiniones delusorias que «los mortales» sostienen acríticamente.

Una de las tareas más antiguas, mas intersantes y, quizá, más importantes de la filosofía es el examen crítico de tales «posiciones» y de las teorías o concepciones que entrañan, sobre todo las que se dan por sentadas. Al hacerlo se descubre con frecuencia que hay *grupos de concepciones relacionadas,* relacionadas por supuestos comunes, por preferencias comunes o por aversiones comunes. Es obvio que a veces es conveniente, o

incluso necesario, dar nombres a esas posiciones o concepciones o grupos de concepciones. De ahí los «ismos».

Que los «ismos» se hayan pasado de moda en la filosofía moderna y que usarlos, o usar nombres similares, se considere señal de mal gusto, es debido al hecho de que la discusión crítica de teorías o posiciones o grupos de concepciones se ha pasado de moda. Pero las modas, sobre todo en filosofía, no deberían aceptarse. Deberían examinarse críticamente, porque ellas mismas no son más que «ismos»: «ismos» acríticamente aceptados.

Todo esto concierne a la inducción. Muchos filósofos e incluso algunos científicos creen que la inducción es un hecho innegable de sentido común: que el uso real de lo que se llama ahora «procedimientos inductivos» no puede rechazarse en serio. Esto puede ser así o no (véase la sección 3, más adelante). Pero, en cualquier caso, deberíamos aprender a escuchar a los que rechazan hechos de sentido común. Los filósofos han tenido una enorme paciencia al escuchar a gente que afirmaba y a otros que negaban, la existencia de mesas, sillas, meteoritos o fantasmas (tanto dentro como fuera de máquinas), o de enunciados analíticos. Pero la discusión de todas esas afirmaciones e «ismos» no es tan fundamental para la filosofía como la discusión sobre la existencia o inexistencia de procedimientos inductivos. Así pues, discutamos críticamente la cuestión. Quizá el supuesto de que los procedimientos inductivos —una «lógica inductiva»— existen es, después de todo, un prejuicio, de modo que todo lo que existe es simplemente un mito, un «ismo» equivocado («inductivismo»).

Si existe algo que podamos llamar inductivismo, entonces esto explica por qué los «ismos» no están de moda. Porque cosa como «ismos» aceptados acríticamente son un peligro para el inductivismo. Bacon lo vio. Pero su remedio —purga tu mente— era ingenuo. De modo que les ha parecido mejor a los inductivistas mirar hacia otro lado o, si no, estudiar los «ismos» inductivamente.

Estas teorías que, si se tienen inconscientemente, se tienen, obviamente, de un modo acrítico, a menudo se incorporan a nuestra lengua; y no sólo a su vocubulario sino también a su estructura gramatical. El primero en ver esto fue, que yo sepa, Bertand Russell, cuando señaló que muchas teorías filosóficas dependen del supuesto erróneo de que «todas las proposiciones son reducibles a la forma de sujeto-predicado», un supuesto que está estrechamente vinculado a la estructura gramatical de las lenguas indoeuropeas[2]. Posteriormente, Benjamín Lee

[2] Bertrand Russel, *A Critical Examination of the Philosophy of Leibniz*,

Whorf desarrolló una doctrina similar y recalcó, sobre todo, la dependencia de nuestra idea del tiempo respecto a nuestro lenguaje[3].

Estos hechos a propósito de la lengua se usan a veces para defender la siguiente conclusión radical: puede decirse que, intelectualmente, todos somos prisioneros de nuestra lengua: no podemos pensar más que en términos de teorías (de sustancia, de espacio y tiempo, por ejemplo) que, sin saberlo nosotros, se incorporan a nuestra lengua; y no podemos escapar por nuestros propios esfuerzos —por ejemplo, por medio de una discusión crítica— de nuestra prisión, porque la discusión crítica tendría que realizarse con ayuda de nuestro lenguaje y, por tanto, permanecería dentro de él, dentro la de la prisión. Sólo aprendiendo una nueva lengua con una estructura diferente —una que, esencialmente, no fuera plenamente traducible a nuestra antigua lengua— y, por tanto, sólo mediante el choque con una nueva cultura, y la conversión a ella, podríamos liberarnos de nuestra prisión; o, más bien, podríamos entrar en otra prisión, posiblemente una mayor.

Me parece a mí que hay mucho de cierto en esta doctrina de encarcelamiento, pero que sus consecuencias se han exagerado. Aunque la ayuda que presta un choque cultural puede tener un valor inmenso, a veces podemos pasarnos sin él: podemos conseguir nuestro objetivo de demoler uno u otro de los muros de nuestra prisión por medio de nuestros esfuerzos críticos. Russel lo hizo cuando descubrió o se hizo consciente de uno de esos muros por medio del estudio crítico de la filosofía de Leibniz. La conciencia de esos muros equivale a su destrucción, puesto que el encarcelamiento es intelectual: consiste, en gran medida, en nuestra ceguera intelectual ante los muros de la prisión.

Hay que admitir que el descubrimiento de Russell y la resultante destrucción de uno de los muros de la prisión no le liberó —ni a nosotros— totalmente. Algunos de los viejos muros siguen en pie, precisamente porque no nos hemos percatado de ellos; e incluso la destrucción del muro intangible, del que era responsable nuestra gramática de sujeto-predicado, no signi-

---

1900. Véase por ejemplo p. 15, donde Russell, supongo que por primera vez en sus escritos, se refiere a los que ha llamado posteriormente «pseudo-proposiciones», es decir, «una forma de palabras sin sentido» que puede confundirse con una proposición. Véase también el artículo de Russell «Logical Atomism» en J. H. Muirhead, ed. *Contemporary British Philosophy*, Primera Serie, 1924, pp. 360 y 367 s., con respecto a «la influencia de la lengua en la filosofía» y, especialmente, sobre «la sustancia-atributo metafísico»; y véase también la p. 61 de su libro *My Philosophical Development*, 1959.

[3] Véase B. L. Whorf, *Language, Thought and Reality*, 1956.

fica que ya podamos escaparnos a terreno abierto. Sólo podemos escaparnos a una prisión más amplia (la de un lenguaje de relaciones). Este hecho, sin embargo, no debe deprimirnos. Una cadena perpetua que nos condene a una prisión intelectual de la que, en principio, podemos librarnos escapando a otra más amplia y luego a otra más amplia aún, sin límites predeterminados, no sólo es una sentencia soportable, sino una sentencia que además abre una emocionante perspectiva de luchar por la libertad: una tarea digna para nuestra vida intelectual.

Hay varias maneras totalmente análogas de estar intelectualmente encarcelado. Podemos estar encarcelados no simplemente en un lenguaje, sino también en diversos sistemas de supuestos o teorías o puntos de vista (que han sido denominados «ideologías totales»[4]) dentro de un lenguaje; supuestos de los que podemos no ser conscientes y que por esta razón podemos ser incapaces de criticar o de transcender. Aunque hay que admitir todo esto, no hay que exagerarlo. Por ejemplo, se dice con frecuencia que una discusión no puede ser fructífera más que si los participantes están de acuerdo en lo fundamental o comparten una formación común o un «marco conceptual» común[5]. Yo niego que tal aserción sea cierta. Aunque una discusión puede ser muy satisfactoria para los participantes si están de acuerdo en todos los puntos principales, será mucho más fructífera en condiciones menos agradables.

Una de las tareas de la crítica filosófica es hacer conscientes esos diversos sistemas de creencias de modo que, tras un examen inquisitivo, podamos escoger tentativamente los mejores de que dispongamos. Pero esto significa que debemos entender, examinar, comparar y criticar sistemas coherentes de supuestos, es decir, «enfoques» o «ismos» (tales como el inductivismo o el positivismo o el intuicionismo). En aquella reunión de la Sociedad Aristotélica, mi intento de explicar mi solución estaba destinado al fracaso porque no tenía oportunidad de analizar y de criticar el enfoque o actitud que he llamado «inductivismo», una actitud que iba a dar a mis palabras una interpretación que yo no me había propuesto.

La gente que piensa tiende a desarrollar un marco en el que tratan de encajar cualquier idea nueva que les salga al paso; por regla general, llegan incluso a traducir cualquier idea nueva

[4] Una crítica de la doctrina de las «ideologías totales» y de similares construcciones sociológicas se encuentra en el capítulo 23 de mi *Sociedad abierta (O. S.), op. cit.*; véase por ejemplo, la referencia a la crítica de Einstein de nuestro «aparato categórico».

[5] [Véase «The Myth of the Framework», *op. cit.*, y «Addendum: Facts, Standards and Truth: A Further Criticism of Relativism» en *La sociedad abierta (O. S.), op. cit., Ed.*]

que encuentran a un lenguaje adecuado a su propio marco. Una de las tareas más características de la filosofía es atacar, si es necesario, al propio marco. Y para hacerlo, puede resultar necesario atacar creencias que, tanto si se tienen de modo consciente como si no, se dan por sentadas hasta tal punto que cualquier crítica que se les haga se considera maligna o insincera. Siempre que se ataque al propio marco, sus defensores interpretarán el ataque y tratarán de refutarlo dentro de su propio marco adoptado. Pero al tratar de traducir los argumentos críticos dirigidos contra el marco a un lenguaje adecuado al marco, corren el riesgo de dar lugar a distorsiones y malentendidos fatales. Una discusión en términos de «ismos» puede disminuir este peligro hasta cierto punto, al resaltar constantemente el hecho de que es al propio marco al que se dirige la acometida.

2. EL ENFOQUE CRITICO: LA SOLUCION DEL PROBLEMA DE LA INDUCCION

Yo no creo en la Creencia.

E. M. Forster

I

Durante los muchos años transcurridos desde aquella reunión en la Sociedad Aristotélica nunca se me había ocurrido que mi propio enfoque, mi propio marco, podía estar más expuesto a malentendidos que otros —que podía chocar más seriamente que otros con ciertas concepciones ampliamente sostenidas e inconscientemente aceptadas— y que como consecuencia, la gente podía malinterpretar mi enfoque identificándolo con alguna forma de irracionalismo, escepticismo o relativismo. Sólo recientemente empecé a sospecharlo y a sospechar que mi teoría del conocimiento era más revolucionaria, y por esa razón más difícil de captar, de lo que yo había pensado. Esta sospecha nació de una nueva forma de concebir mi propio enfoque y su relación con la situación de los problemas en filosofía; una forma que me fue sugerida por mi amigo W. W. Bartley III. Sus concepciones son impresionantes en sí mismas[6]. Pero también explican por qué ciertos malentendidos de mi posición surgen casi por necesidad.

[6] Algunas de estas concepciones han sido publicadas; véase W. W. Bart-

El problema central de la filosofía del conocimiento, al menos desde la Reforma, ha sido éste: ¿Cómo podemos juzgar o valorar las afirmaciones de largo alcance de teorías y creencias rivales? Llamaré a éste nuestro *primer* problema. Este problema ha llevado, históricamente, al *segundo* problema: ¿Cómo podemos *justificar* nuestras teorías o creencias? Y este segundo problema está, a su vez, ligado a otra serie de interrogantes: ¿En qué consiste una justificación? Y más especialmente: ¿Es posible justificar nuestrar teorías o creencias *racionalmente*, esto es, dando razones, «razones positivas» (como las llamaré), tales como la apelación a la observación; es decir, razones para sostener que son verdaderas o que son, por lo menos, «probables» (en el sentido del cálculo de probabilidades)? Claramente, aquí hay un supuesto no enunciado y aparentemente inocuo que patrocina la transición de la primera cuestión a la segunda: a saber, que uno juzga entre aserciones rivales determinando cuál de ellas puede ser *justificada* por medio de razones positivas y cuál no.

Ahora bien, Bartley sugiere que mi enfoque resuelve el primer problema, pero que al hacerlo, cambia por completo su estructura. Porque yo rechazo el segundo problema por no pertinente y las respuestas que se le dan habitualmente, por incorrectas. También rachazo por incorrecto el supuesto que lleva del primer problema al segundo. Afirmo (a diferencia, sostiene Bartley, de todos los racionalistas anteriores, excepto, quizá, los que fueron arrastrados al escepticismo) que no podemos dar ninguna justificación positiva o ninguna razón positiva de nuestras teorías o de nuestras creencias. Es decir, que no podemos dar ninguna razón positiva para sostener que nuestras teorías son *verdaderas*. Afirmo, además, que la creencia de que podemos dar tales razones y que debemos buscarlas no es, en sí misma, ni una creencia racional ni una creencia verdadera, sino una de la que puede mostrarse que carece de méritos.

(Estaba a punto de escribir «infundada» donde he escrito que «carece de méritos». Esto es un buen ejemplo de lo influido que está nuestro lenguaje por los supuestos inconscientes que mi propio enfoque ataca. Se supone, sin crítica, que sólo una concepción que carece de mérito tiene que ser infundada —sin fundamento, en el sentido de sin base, o injustificada o carente

---

ley III, *The Retreat To Commitment*, 1962, y «Rationality versus the Theory of Rationality», en Mario Bunge, ed., *The Critical Approach to Science and Philosophy*, 1964. [Ahora, véase también mi «The Philosophy of Karl Popper, Part III: Rationality, Criticism, and Logic», *Philosophia*, 1982, *Ed.*] Mis comentarios en el texto no se basan en estas publicaciones, sino en conversaciones con Bartley, antes de su publicación. La presente sección se rehizo en parte en 1979.

de sostén—. Mientras que en mi concepción, *todas* las concepciones —buenas y malas— son, en este importante sentido, carentes de base, infundadas, injustificadas, carentes de sostén.)

En la medida en que mi enfoque entraña todo esto, mi solución al problema central de la justificación —tal como se ha entendido siempre— es tan *inequívocamente negativa* como la de cualquier irracionalista o escéptico.

Con todo, difiero del escéptico y del irracionalista en que ofrezco una solución *inequívocamente afirmativa* a otro problema, el tercero, que, aunque es similar al problema de si podemos o no dar razones positivas válidas para sostener que una teoría es *verdadera*, debe distinguirse estrictamente de él. Este tercer problema es el problema de si una teoría es *preferible* a otra y, si es así, por qué. (Estoy refiriéndome a que una teoría sea preferible en el sentido de que pensamos o conjeturamos que es una *aproximación más cercana a la verdad* y que incluso tenemos *razones* para pensar o conjeturar que lo es.)

Mi respuesta a esta cuestión es inequívocamente afirmativa. A menudo podemos dar razones para considerar una teoría preferible a otra. Estas consisten en señalar que una teoría ha resistido las críticas mejor que otra, hasta el momento, y cómo lo ha hecho. Llamaré a esas razones *razones críticas*, para distinguirlas de las *razones positivas* que se ofrecen con la intención de justificar una teoría o, en otras palabras, de justificar la creencia en su verdad.

Las razones críticas no justifican una teoría, porque el hecho de que una teoría haya resistido a la crítica mejor que otra no es razón, en modo alguno, para creer que es realmente verdadera. Pero, aunque las razones críticas no pueden justificar una teoría, pueden usarse para defender (pero no para *justificar*) nuestra *preferencia* por ella: es decir, nuestra decisión de usar ésa, en vez de algunas o todas las teorías propuestas hasta el momento. Tales razones críticas no prueban, desde luego, que nuestra preferencia sea más que conjetural: tendremos que renunciar a nuestra preferencia si nuevas razones críticas hablan en contra de ella o si se propone una nueva y prometedora teoría, lo que exige la reapertura de la discusión crítica.

El dar razones en favor de las propias preferencias puede *llamarse* justificación, naturalmente (en el lenguaje corriente). Pero no es una justificación en el sentido que aquí se critica. Nuestras preferencias están «justificadas» sólo en relación con el estado actual de nuestra discusión.

Dejando para más tarde la importante cuestión de las normas de preferencia para las teorías, presentaré ahora la concepción de Bartley sobre la nueva situación del problema que ha

surgido. Bartley describe la situación de forma extraordinaria diciendo que, después de haber dado una solución negativa al *problema de la justificación*, lo he sustituido por el *problema de la crítica*, un problema nuevo para el que ofrezco una solución afirmativa.

Esta transición del problema de la justificación al problema de la crítica es fundamental, sugiere Bartley, y da lugar a malentendidos porque casi todo el mundo da implícitamente por sentado que todos los demás (incluido yo) aceptan el problema de la justificación como el problema central de la teoría del conocimiento.

Porque, según Bartley, *todas las filosofías, hasta el presente, han sido filosofías justificacionistas*, en el sentido de que todas suponían que era la tarea *prima facie* de la teoría del conocimiento mostrar que podemos *justificar* nuestras teorías o creencias y cómo. No sólo los racionalistas, los empiristas y los kantianos compartían este supuesto, sino también los escépticos y los irracionalistas. Los escépticos, obligados a admitir que no podemos justificar nuestras teorías o creencias, declaran la bancarrota de la búsqueda del conocimiento; mientras que los irracionalistas (por ejemplo, los fideístas), a causa de la misma admisión fundamental, declaran la bancarrota de la búsqueda de razones —es decir, de argumentos racionalmente válidos— y tratan de justificar nuestro conocimiento, o más bien nuestras creencias, recurriendo a la autoridad, tal como la autoridad de las fuentes irracionales. Los dos suponen que la cuestión de la justificación, o de la existencia de razones positivas, es fundamental: los dos son justificacionistas clásicos.

Bartley observa que mi enfoque ha sido confundido normalmente con una forma de justificacionismo, aunque, en realidad, es totalmente diferente de él. Porque, aunque ofrezco una solución negativa al problema clásico de la justificación, y en este aspecto me parezco a los escépticos y a los irracionalistas, al mismo tiempo, destrono al problema clásico y lo sustituyo por un nuevo problema central que admite una solución que no es ni escéptica ni irracionalista. Porque la solución que yo propongo para el nuevo problema es compatible con la concepción de que nuestro conocimiento —nuestro conocimiento conjetural— puede crecer y que puede hacerlo con el uso de la razón, del argumento crítico.

Mi posición, sugiere Bartley, corre el riesgo de ser mal entendida, a menos que se comprenda primero que el problema clásico de la justificación ha sido no sólo desplazado de su posición central, sino que, visto desde el nuevo punto de vista, debe descartarse como insignificante. Sin embargo esto es difícil de ver para el justificacionista. Porque él arguye como Hume:

«Si yo os pregunto por qué creéis en una determinada cuestión de hecho... debéis darme alguna razón; ... o debéis admitir que vuestra creencia carece por completo de fundamento»[7].

Ahora bien, igual que E. M. Forster, yo no creo en la creencia: no estoy interesado en una filosofía de la creencia y no creo que las creencias y su justificación o fundamento o racionalidad sean el objeto de la teoría del conocimiento. Pero si en este pasaje de Hume sustituimos las palabras «creéis en» por «proponéis una teoría o una conjetura sobre» y las palabras «vuestra creencia» por «vuestra conjetura», entonces su declaración pierde fuerza. Porque pocos se extrañarán al oír que su conjetura «carece por completo de fundamento». Tener algún fundamento o justificación puede ser importante para una creencia; pero no es el tipo de cosa que requerimos de una conjetura o hipótesis, al menos no en el sentido en que Hume usa el término «fundamento» (que corresponde a mi frase «justificación por medio de razones positivas»). Hay que admitir que algunos hablan de los «fundamentos de la teoría física», por ejemplo; pero esto, o es charla justificacionista, o significa algo muy diferente; y una vez que nos damos cuenta de que las teorías físicas son conjeturas o hipótesis y que están sujetas a cambios revolucionarios, podemos preferir no hablar de sus «fundamentos», ni más ni menos que sobre nuestra creencia en ellas.

Admitamos que pueden darse algunas razones para *proponer* una hipótesis y para someterla a discusión crítica. Pero ésas no son razones justificativas sino que tienen, más bien, la naturaleza de las *explicaciones* de por qué —*a la luz de nuestros objetivos,* tales como conseguir teorías más criticables y más severamente criticadas— ofrecemos una teoría en vez de otra. Esas razones y su papel lógico son enteramente diferentes de las que Hume tenía en mente. Podemos ofrecer, por ejemplo, una razón perfectamente buena para proponer una hipótesis, señalando que, si fuera verdadera, resolvería un problema que queremos resolver (como la teoría de Newton, que resolvió el problema de explicar las leyes de Kepler). Una razón de este tipo puede ser suficiente para proponer una hipótesis y recomendarla como digna de nuestra atención crítica. Pero, naturalmente, no sería una razón para suponerla verdadera. Puede que no sea siquiera una razón para aceptarla tentativamente, o incluso para preferirla, porque puede haber otras hipótesis conocidas que resuelvan el problema aún mejor.

---

[7] David Hume, *Enquiry Concerning Human Understanding,* Sección V, Parte I; Selby-Bigge, p. 46. [*Cf.* Popper sobre métodos de crítica en *El yo y su cerebro (S. & I. B.), op. cit.,* pp. 172-73. *Ed.*]

II

De este modo nos damos cuenta de que el problema epistemológico de Hume —el problema de dar razones positivas, justificativas, o el problema de la justificación— puede ser reemplazado por el problema totalmente diferente de explicar —dar razones críticas— de por qué preferimos una teoría a otra (o a todas las demás que conocemos) y finalmente por el problema de *la discusión crítica* de las hipótesis para descubrir cuál de ellas es, comparativamente, la que hay que preferir.

Un justificacionista puede objetar, sin embargo, que no he sustituido realmente un problema por otro. Puede argüir, en primer lugar, que, en vez de «razones por las que preferimos una teoría a otra», yo podría haber dicho «razones por las que *creemos* que una teoría es mejor que otra». En la medida en que este argumento es verbal, estoy dispuesto a conceder que es así; porque aunque no quiero filosofar sobre creencias, nunca discuto sobre las palabras. En segundo lugar, puede señalar que, incluso aunque él admita que quizá esas «razones por la que creemos que una teoría es mejor que otra» no tengan el mismo carácter que tendrían las razones para creer, por ejemplo, que la primera de esas teorías es *verdadera*, todavía podrá mantener que son «razones positivas»: que son razones para creer en la verdad de *una* teoría, es decir la teoría (la «meta-teoría», como podría llamársela) de que la primera teoría es mejor que la segunda. De esta manera, el justificacionista podría llegar a la conclusión de que realmente no he sustituido el problema de la justificación por uno diferente.

Sin embargo, al decir esto, el justificacionista simplemente no se daría cuenta de lo que ha admitido. Primero, hay una diferencia inmensa entre una meta-teoría que afirma que una teoría *A* es, de hecho, mejor que una teoría *B*, y otra meta-teoría que afirma que la teoría *A* es, de hecho, verdadera (o «probable»). Y hay una diferencia inmensa entre los argumentos que podrían considerarse como razones válidas o razones de peso en apoyo de una o la otra de esas dos meta-teorías. Por ejemplo, al discutir teorías o conjeturas explicativas rivales (por ejemplo, sobre la estructura de la materia), a menudo podemos resumir la situación con imparcialidad diciendo que, de acuerdo con el estado actual del debate crítico, la conjetura *a* es muy superior a la conjetura *b* o incluso a todas las demás conjeturas propuestas hasta el momento: que parece ser una aproximación a la verdad mejor que cualquiera de ellas (y quizá que *puede* incluso ser realmente verdadera). Pero, en general, no seremos capaces de decir que, conforme al estado

actual del debate crítico, la conjetura *a* es la mejor que presentará nunca en este terreno o que *parece* ser realmente verdadera. Así, una de las dos meta-teorías puede no hacer más que resumir el estado actual de nuestra *discusión crítica* de modo imparcial, y puede, en ese sentido, ser meramente negativa, meramente crítica; mientras que la otra, en general, no hará nada de esto en absoluto (aunque es posible que resuma el estado actual de nuestra creencia o de nuestra convicción intuitiva).

En segundo lugar, no hay, una vez más, ninguna pretensión por mi parte de *justificar* positivamente o de establecer, en el sentido tradicional, que la preferencia por una teoría en lugar de por otra, es la correcta. No se ha cambiado simplemente el problema de la justificación: se ha suprimido. Tampoco la meta-teoría se justifica positivamente; es conjetural y abierta a la crítica.

III

Una objeción más importante parece ser la siguiente. Mi afirmación de haber sustituido el problema de la justificación por otro es tan infundada, podría decir el justificacionista, como mi afirmación de haber dado soluciones que difieren de las de los escépticos e irracionalistas. Que la última afirmación es infundada puede verse en el hecho, podría argüir el justificacionista, de que mi respuesta ser idéntica o, en el mejor de los casos, una variante del relativismo, pragmatismo y otras concepciones igualmente conocidas. Porque al decir que debemos sustituir la cuestión de si una teoría es verdadera por la cuestión de si es mejor o peor que otra teoría, yo adopto claramente, puede argüir el justificacionista, una posición relativista con respecto a la verdad. Y cuando digo que la última cuestión debe decidirse recurriendo al éxito de esas teorías, me revelo como un pragmatista (puede argüir) o quizá incluso como un convencionalista.

Pero esta objeción me atribuye implícitamente doctrinas justificacionistas que yo no sostengo. Porque no digo que debamos *sustituir* la cuestión de si una teoría es verdadera por la cuestión de si es mejor que otra; y tampoco digo que una teoría sea mejor que otra siempre que tenga más éxito en un sentido pragmático. Ambos puntos son de gran importancia.

Mi posición es ésta. Yo afirmo que la búsqueda de la verdad —o de una teoría verdadera que pueda resolver nuestro problema— es de la mayor importancia: *toda crítica racional es la crítica a la afirmación de que una teoría es verdadera y capaz*

de resolver los problemas para cuya solución se concibió. Por tanto, no *sustituyo* la cuestión de si una teoría es verdadera por la cuestión de si es mejor que otra. Más bien, sustituyo la cuestión de si podemos presentar *razones* válidas (razones positivas) en favor de la verdad de una teoría por la cuestión de si podemos presentar *razones* válidas (razones críticas) contra su afirmación de que es verdadera o contra la verdad de sus rivales. Además, describir una teoría como mejor que otra, o superior, o lo que sea, equivale a indicar, en mi concepción, que parece *más cercana a la verdad*[8].

La verdad —la verdad absoluta— continúa siendo nuestro objetivo; y continúa siendo el criterio implícito de nuestra crítica: casi toda crítica es un intento de refutar la teoría criticada: es decir, de mostrar que *no es verdadera.* (Una excepción importante es una crítica que intente mostrar que una teoría no es pertinente, que no resuelve el problema para cuya solución se concibió.) Por tanto, siempre estamos buscando una *teoría verdadera* (una teoría verdadera y pertinente), aunque nunca podamos dar razones (razones positivas) para mostrar que hemos encontrado realmente la teoría verdadera que habíamos estado buscando. Al mismo tiempo, podemos tener buenas razones —es decir, buenas *razones críticas*— para pensar que hemos aprendido algo importante: que hemos avanzado hacia la verdad. Porque, en primer lugar, podemos haber aprendido que una determinada teoría no es verdadera según el estado actual de la discusión crítica; y, en segundo lugar, podemos haber encontrado algunas razones provisorias para creer (sí, incluso para creer) que una nueva teoría se acerca más a la verdad que sus predecesoras.

Para ser menos abstracto, daré un ejemplo histórico.

Las teorías de Einstein han sido muy discutidas por los filósofos, pero pocos de ellos han resaltado el importante hecho de que Einstein no creía que la relatividad especial fuera verdadera: desde el principio señaló que, en el mejor de los casos, podía ser una aproximación (ya que sólo era válida para el movimiento no acelerado)[9]. De modo que avanzó hacia una mayor aproximación, la relatividad general. Y nuevamente señaló que su teoría no podía ser verdadera tampoco, sino sólo

---

[8] Para un análisis lógico del término «más cercana a la verdad», véase *El desarrollo del conocimiento científico: conjeturas y refutaciones (C. & R.,* 1963) sobre todo capítulo 10 y el Addendum. Véase también la sección 4 más abajo. [Véase también *Conocimiento objetivo (Obj. Kn.,* 1972), especialmente capítulos 2, 3 y 9; y *El yo y su cerebro (S. & I. B.,* 1977), pp. 148-49. *Ed.*]

[9] [Véase el «Prefacio de 1982» de Popper a *Teoría cuántica y el cisma en física (Q. T. & S. Ph.)* Tomo III del *Post Scriptum. Ed.*]

una aproximación. De hecho, buscó una aproximación mejor durante casi cuarenta años hasta su muerte[10].

No hay indicio de relativismo epistemológico en la actitud de Einstein, a pesar del nombre «Teoría de la Relatividad»: buscó la verdad y pensó que tenía razones —razones críticas— que indicaban que no la había encontrado. Al mismo tiempo, él (y otros muchos) dieron razones críticas de que había avanzado enormemente hacia ella, que sus teorías resolvían problemas que sus predecesoras no podían resolver, que se acercaban más a la verdad que sus rivales conocidas[11].

Este ejemplo puede apoyar mi afirmación de que al sustituir el problema de la justificación por el de la crítica no necesitamos renunciar ni a la teoría clásica de la verdad como correspondencia con los hechos ni a la aceptación de la verdad como uno de nuestros criterios para la crítica. (Otros principios son pertinencia respecto a nuestros problemas y capacidad explicativa.)

Así, aunque yo sostengo que la mayoría de las veces no encontramos la verdad y que, incluso cuando la hemos encontrado, no lo sabemos, retengo la idea clásica de la verdad absoluta u objetiva como *idea regulativa:* es decir, *como un criterio respecto al cual nos quedamos cortos.* El cambio no se ha hecho con respecto a la idea de la verdad, sino con respecto a cualquier pretensión de que conocemos la verdad; esto es, de que disponemos de argumentos o razones que son suficientes, o incluso que son casi suficientes, para establecer la verdad de cualquier teoría de que se trate.

No hay necesidad de extrañarse ante el descubrimiento de que no podemos justificar ni siquiera apoyar con argumentos o

---

[10] Tal como se vio despues de su muerte, es posible que Einstein estuviera trágicamente equivocado en sus dudas respecto a la relatividad general; véanse, por ejemplo, Charles W. Misner y John A. Wheeler, «Classical Physics as Geometry», *Annals of Physics* 2, 1957, pp. 525 ss., donde se llama la atención sobre G. Y. Rainich, *Transactions of the American Mathematical Society* 27, 1925, pp. 106-136. Véase también Rainich, «Electromagnetics in the General Relativity Theory», *Proceedings of the National Academy of Sciencies of the U. S. A.* 10, 1924, pp. 124 ss.; *The Mathematics of Relativity,* Nueva York, 1950; y los últimos artículos de Wheeler, sobre todo su informe en E. Nagel, P. Suppes y A. Tarski, eds. *Logic, Methodology, and Philosophy of Science, Proceedings of the 1960 International Congress,* 1962, pp. 361-374. *Desde que se escribió esta nota, la teoría de Misner y Wheeler, llamada ahora «geometrodinámica» se ha desarrollado enormemente. Véase sobre todo Wheeler, *Geometrodinamics,* 1962; y Charles W. Misner, Kip S. Thorne y John A. Wheeler, *Gravitation,* 1973.

[11] Las razones de Einstein para pensar así fueron criticadas de un modo muy interesante por Alfred O'Rahilly, en su *Electromagnetics: A Discusion of Fundamentals,* 1938, un libro que destaca entre los escritos dedicados al ataque (y habitualmente a la tergiversación) de la relatividad especial.

razones la afirmación de que nuestras teorías son verdaderas. Porque el razonamiento crítico tiene todavía una función de la mayor importancia con respecto a la evaluación de las teorías: podemos criticar y distinguir entre nuestras teorías como consecuencia de nuestra discusión crítica. Aunque, por regla general, en dicha discusión no podemos distinguir (con certeza o casi certeza) entre una teoría verdadera y una falsa, a veces podemos distinguir entre una teoría y una que *puede* ser verdadera. Y, a menudo, podemos decir de una teoría determinada que, a la luz del estado actual de nuestra discusión crítica, parece ser mucho mejor que cualquier otra de las teorías presentadas; mejor, naturalmente, desde el punto de vista de nuestro interés por la verdad; o mejor en el sentido de un mayor acercamiento a la verdad.

Así, subrayo la función crítica (o, si lo prefieren, negativa) de la razón. Sin embargo, también subrayo que razonar es más importante, más eficaz y menos estéril de lo que normalmente se piensa. La crítica racional es, desde luego, el medio por el que aprendemos, aumentamos nuestro conocimiento y nos transcendemos a nosotros mismos.

IV

Me parece a mí que la sencilla formulación de Bartley —que la *justificación* puede sustituirse *por crítica no justificadora*— y su énfasis en el cambio de foco que implica la transición desde las diversas *filosofías justificacionistas* a una *filosofía crítica que no aspira a la justificación* es muy esclarecedora; al menos yo la considero así; y, bajo la impresión de que el lector podría beneficiarse también de ella, he decidido presentarla aquí. Me ha ayudado a ver por qué la idea misma de crítica es, tan frecuentemente, mal entendida por los filósofos justificacionistas: ellos tiende a reducir la idea de crítica válida a la limitada tarea de demostrar la falta de validez de ciertos intentos de *justificar* ciertas creencias. La formulación de Bartley también ayuda a explicar por qué puedo yo estar de acuerdo con tanto de lo que han dicho diversos irracionalistas contra el racionalismo y los intentos racionalistas de justificar nuestras creencias, a pesar de que no haga concesiones al irracionalismo, sino que, por el contrario, insisto en que cualquier teoría o creencia puede y debe ser sometida a una crítica *racional* severa e inquisitiva y que debemos buscar razones —argumentos *racionales*— que pudieran refutarla. En realidad he sugerido

que lo que distingue la actitud de un racionalista es sencillamente su apertura a la crítica[12].

Los irracionalistas tienen bastante razón cuando insisten en que tenemos «fuentes de conocimiento» distintas de la razón y la observación; por ejemplo la inspiración o la compasión; o la tradición, que es quizá la «fuente de conocimiento» más importante y que con tanta frecuencia es ignorada por los racionalistas a causa de su falibilidad[13]. Pero los irracionalistas se equivocan peligrosamente cuando sugieren que hay algún conocimiento, del tipo o la fuente o el origen que sea, que está por encima de la crítica racional o exento de ella.

V

Igualmente equivocada es la concepción de que la crítica racional corre la misma suerte que el argumento racional positivo, puesto que también ella debe basarse siempre en algunas presuposiciones no demostrables, de forma que su validez es esencialmente *relativa* a esas presuposiciones; o, en otras palabras, que nos enfrentamos a una situación en la que la *A* se critica en términos de la posición *B*, que, sin embargo, es a su vez imposible de establecer; de modo que ninguna crítica de *A* en términos de *B* será concluyente, y viceversa, naturalmente.

La mayor parte de los componentes individuales de este argumento son bastante correctos. Pero la conclusión no se desprende de ellos en absoluto. Así pues, estoy dispuesto a admitir que, en nuestras críticas, a menudo operamos con presuposiciones injustificables y no demostrables. Y, por tanto, nuestras críticas no son nunca concluyentes. Pero la indemostrabilidad de cualquier tipo nunca preocupa al crítico racionalista. Porque sus argumentos críticos —de la misma manera que las teorías que está criticando por medio de ellos— son conjeturales. La diferencia es muy sencilla. El argumento justificador, que retrocede a razones positivas, acaba llegando a razones que no pueden ser justificadas (si no fuera así, el argumento llevaría a una retrogresión infinita). Y el justificacionista concluye, normalmente, que esas «presuposiciones

[12] Véase por ejemplo mi libro *La sociedad abierta y sus enemigos* (*O. S.*, cuarta edición inglesa, 1962) capítulo 24; y mi conferencia en la Academia, «Sobre las fuentes del conocimiento y la ignorancia» (que es ahora la Introducción a *El desarrollo del conocimiento científico* [*C. & R.*, 1963]). [Véase también el «Addendum: Facts, Standards, and Truth: A Further Criticism of Relativism», *op.cit. Ed.*]

[13] [Véase Popper, «Hacia una teoría racional de la tradición», *El desarrollo del conocimiento científico (C. & R.) Ed.*]

últimas» deben estar, en algún sentido, más allá del argumento y no pueden ser criticadas. Pero las críticas, las razones críticas, que se proponen en mi enfoque no son en *ningún* sentido últimas; *también ellas están abiertas a la críticas;* son conjeturales. Uno puede seguir indefinidamente examinándolas; están indefinidamente abiertas al reexamen y a la reconsideración. Y, no obstante, no se produce una *retrogresión infinita:* porque no se trata de probar o establecer o justificar nada; y no hay necesidad de ninguna presuposición *última.* Sólo la exigencia de prueba o de justificación produce una retrogresión infinita, porque crea la necesidad de un *término* último de la discusión. Este es el fondo de la diferencia entre justificación y crítica[14].

Relacionada con esta objeción está la concepción, ampliamente sostenida, de que un método puramente crítico —es decir, uno que se abstenga de justificar nada positivamente— es imposible, ya que tendría que reducirse a una crítica del tipo «inmanente». Una crítica se llama «crítica inmanente» si ataca una teoría desde dentro, adoptando todos sus supuestos y presuposiciones y sólo éstos; y se llama «crítica transcendente» si ataca a una teoría desde fuera, procediendo a partir de supuestos o presuposiciones que son ajenas a la teoría criticada. Sin embargo, esto es lo que se dice, la crítica inmanente es relativamente poco importante; porque, como no puede hacer más que señalar las inconsistencias lógicas *dentro* de la teoría criticada, nunca tiene éxito contra una teoría consistente. (Podría incluso decirse, además, que el señalar inconsistencias puede no ser totalmente inmanente, ya que supone o presupone una lógica: una que prohíbe las inconsistencias.) De este modo se llega a la conclusión de que toda crítica a una teoría consistente, y así, la crítica más importante, tiene que ser «transcendente». De ahí que se sostenga que los defensores de una teoría que está siendo atacada siempre pueden rechazar toda crítica por no concluyente o no válida, a menos, naturalmente, que los supuestos o presuposiciones que fundamenten una crítica transcendente puedan recibir una *justificación positiva.* Así, los métodos de la crítica y de la justificación positiva corren, al parecer, la misma suerte.

Esta plausible concepción de la reducida extensión y relativa insignificancia del método es equivocada. Parece estar relacionada con la observación correcta de que todos los críticos de un sistema de pensamiento, o la mayor parte de ellos, adoptan otro sistema de pensamiento que los guía en su ataque crítico. Aunque la crítica puede ser, efectivamente, «transcendente» en lo que respecta a su origen o guía o inspiración, esto no signi-

[14] [Véase Popper, *L. I. C.*, sección 29. *Ed.*]

fica que tenga que ser «transcendente» en el sentido lógico del término. De hecho, ningún crítico que sea auto-crítico estará, por lo general, satisfecho de su crítica a menos que pueda liberarse de todo vestigio de su origen transcendente: aunque quizá, guiado por su propio sistema de pensamiento, transformará su crítica hasta que se convierta en inmanente, y con ello sea más eficaz contra su oponente. Porque la teoría bajo examen no es simplemente un sistema de supuestos, dogmas, conjeturas y demás; es *también un intento de resolver un problema.* Por tanto, *puede* ser criticada inmanentemente, por no resolver sus problemas, por ejemplo, o por no conseguirlo mejor que sus rivales, o por cambiar simplemente el problema a resolver, etc. De este modo, la crítica inmanente puede señalar serias deficiencias incluso en su teoría consistente. Respecto a señalar inconsistencias, esto se aceptará en la mayoría de los casos como un ejercicio de crítica inmanente precisamente porque el problema que trata de resolver una teoría es el de dar una explicación *consistente* de algo (el resultado de un experimento, por ejemplo) que contradecía las teorías anteriores. En general, la situación de los problemas que estimula a la teoría sometida a crítica, y que es lo único que da razón de ser a la teoría, contiene siempre supuestos o presuposiciones que incluyen la aceptación de principios lógicos (y van bastante más allá de ella) —principios como la regla de rechazar las contradicciones— y que proporcionan amplia base para la crítica inmanente. (Por ejemplo, si la situación del problema que lleva a la formulación de una teoría supone la tarea de explicar ciertas observaciones o experimentos, entonces podrán usarse otros experimentos para criticar a la teoría *inmanentemente,* siempre, naturalmente, que los defensores de la teoría estén dispuesto a aceptar los resultados de esos experimentos)[15]. Así pues, la crítica inmanente es posible e importante.

No obstante, no tenemos necesidad de limitarnos a la crítica inmanente. Tenemos perfecto derecho a emplear la crítica transcendente. Esto consiste en partir de una teoría rival; y no hay nada indebido en tratar de mostrar que una teoría presenta deficiencias y otra no. Por el contrario, este tipo de crítica mutua transcendente puede ser la que al final nos permita decir que una teoría es preferible a su rival y por qué.

[15] Debo la observación entre paréntesis a un comentario de Alan E. Musgrave sobre este pasaje.

VI

El enfoque crítico que he descrito aquí lleva casi de inmediato a una solución directa del problema de Hume sobre la inducción (1739)[16].

Recordemos lo que Hume trató de mostrar (acertadamente, en mi opinión, en lo que respecta a la lógica).

i) Indicó que hay innumerables regularidades (aparentes) en la naturaleza sobre las cuales todo el mundo confía en la práctica, y muchas leyes universales de la naturaleza aceptadas por los científicos, que tienen la mayor importancia teórica.

ii) Trató de mostrar que cualquier inferencia inductiva —cualquier razonamiento desde casos singulares y observables (y su ocurrencia repetida) a cualquier cosa como regularidades o leyes— tiene que ser *inválida.* Cualquier inferencia de esa índole, trató de mostrar, no puede ser ni siquiera aproximada o parcialmente válida. No puede ser siquiera una inferencia probable: tiene, por el contrario, que se totalmente infundada y siempre continuará siéndolo, por grande que sea el número de casos observados. Hume trataba de mostrarnos así que no podemos razonar con validez desde lo conocido a lo desconocido, o de lo que se ha experimentado a lo que no se ha experimentado (y así, por ejemplo, del pasado al futuro): por muy frecuentemente que se haya observado salir y ponerse al sol, incluso el mayor número de casos observados, no constituye lo que he llamado una razón positiva en favor de la regularidad o la ley de la salida y la puesta del sol. Así pues, el número de casos observados no puede establecer esta ley ni hacerla probable.

[16] La crítica de la inducción, de Hume, aparece por primera vez en su *Treatise of Human Nature,* Libro I, Parte III, sección VI, donde explica que no debemos, cuando intentamos justificar nuestras creencias, apelar a la experiencia «fuera de aquellos casos particulares que estén incluidos en la observación» (véase p. 91 de la edición de Selby-Bigge); es decir, fuera de lo que sabemos realmente por observación. (Aunque a menudo menciona una inferencia del pasado al futuro, sugiero que esto sólo debemos tomarlo como un caso especial de la inferencia de lo que es realmente conocido por observación a lo que no es conocido.) La discusión sobre el argumento de Hume contra la inducción se continúa en la Parte 2, sección 16 de este tomo, véase la nota 17 a esa sección y también, Hume, *Enquiry Concerning Human Understanding,* Sección IV, Parte II. Para una extensión, del propio Hume, de su crítica a todas las teorías probabilistas de la inducción, véase su obra *An Abastract of a Book Lately Published entitled A Treatise of Human Nature* y mi *L. I. C.,* sección 81 y Apéndice *IX. Desgraciadamente, Hume mezcló su crítica de la inducción con su crítica de la causalidad, de la que, sin embargo, es lógicamente independiente. Con la notabilísima excepción de Bertrand Russel (véase más abajo) los comentaristas no han desenredado las dos cuestiones, por lo general.

iii) Hume señaló que no puede haber razones válidas que justifiquen la creencia en una ley universal, distintas que las proporcionadas por la *experiencia*.

El choque entre i) por un lado, y ii) e iii) por otro, constituye el problema de Hume, *el problema lógico de la inducción*.

Los puntos ii) e iii) pueden reformularse con un poco más de exactitud y brevedad, como sigue:

ii) No puede haber razonamiento válido desde enunciados de observaciones singulares a leyes universales de la naturaleza y, por tanto, a teorías científicas.

Este es *el principio de la no validez de la inducción*.

iii) Exigimos que nuestra adopción y nuestro rechazo de las teorías científicas dependa de los resultados de la observación y el experimento y, por tanto, de enunciados de observaciones singulares.

Este es *el principio del empirismo*[17].

Ahora, demos por sentado i). Entonces, el problema lógico de la inducción consiste en el choque aparente entre ii), el principio de la no validez de la inducción, e iii), el principio del empirismo: el empirismo parece suponer que sin inducción no podemos tener conocimiento científico.

Hume se dio cuenta de que el choque entre ii) e iii) era sólo aparente, porque aceptó ii) e iii), y deshizo el «choque» renunciando al racionalismo. Decidió que todo nuestro conocimiento sobre leyes se obtiene de la observación —de acuerdo con iii)— por inducción y concluyó que, puesto que la inducción no es válida lógicamente, ello muestra que tenemos que basarnos en la asociación («hábito», que resulta de la repetición) en vez de basarnos en la razón.

También yo acepto ii) e iii) pero no saco de ellos ninguna conclusión anti-racionalista. No sólo afirmo la compatibilidad de ii) e iii), sino que afirmo, además, que ii) e iii) son consistentes con el siguiente principio iv):

iv) Exigimos que nuestra adopción y nuestro rechazo de teorías científicas dependan de nuestro *razonamiento crítico* (combinado con los resultados de la observación y el experimento, como lo exige iii)).

Este es *el principio del racionalismo crítico*.

Para ver que i) e iv) son consistentes, no tenemos más que percatarnos de que nuestra «adopción» de teorías científicas sólo puede ser tentativa; que son siempre y siempre serán *suposiciones o conjeturas o hipótesis*. Se exponen, naturalmente, con la esperanza de dar con la verdad a pesar de que lo más frecuente es que yerren. Pueden contrastarse por medio de la

[17] [Véase *Conocimiento objetivo (Obj. Kn.)*, p. 25. *Ed.*]

observación (la tarea principal de la ciencia es hacer esas contrastaciones cada vez más rigurosas), y rechazarse si no resisten la contrastación. Nada en el argumento de Hume habla en contra de la posibilidad de contrastaciones o de rechazar una ley universal porque es contradicha por enunciados de observación. En efecto, con una ley propuesta no podemos hacer más que contrastarla: no sirve de nada pretender que hemos establecido teorías universales o que las hemos justificado o que las hemos hecho probables, mediante la observación. No hemos hecho tal cosa, de ningún modo, y no podemos hacerla. No podemos dar ninguna razón positiva en su favor. Siguen siendo conjeturas o suposiciones, aunque quizá sean unas conjeturas o suposiciones bien contrastadas. Sin embargo, si consideramos los problemas que resuelven y las críticas y las contrastaciones que han resistido, podemos tener excelentes razones críticas para preferirlas a otras teorías, aunque sólo de modo tentativo y provisional.

Lo que he dicho aquí proporciona una solución completa al problema lógico de Hume sobre la inducción. La clave de esta solución es el reconocimiento de que nuestras teorías, incluso las más importantes, e incluso las que son realmente verdaderas, siempre siguen siendo suposiciones o conjeturas. Si son verdaderas de hecho, no podemos conocer este hecho; ni por la experiencia ni por ninguna otra fuente.

Los principales puntos de mi solución son:

i) La aceptación de la concepción según la cual las teorías son de importancia suprema, tanto para la ciencia teórica como para la ciencia práctica.

ii) La aceptación del argumento de Hume contra la inducción: cualquier esperanza de que podamos tener razones positivas para creer en nuestras teorías queda destruida por ese argumento. (Pero recuérdese que el argumento de Hume no presenta ninguna dificultad a los que sostienen que podemos contrastar nuestras teorías en un intento de refutarlas.)

iii) La aceptación del principio del empirismo: las teorías científicas son rechazadas o adoptadas (aunque sólo temporal y tentativamente) a la luz de los resultados de contrastaciones experimentales o basadas en la observación.

iv) La aceptación del racionalismo crítico: las teorías científicas son rechazadas o aceptadas (aunque sólo temporal y tentativamente) por ser peores o mejores que otras teorías conocidas, a la luz de los resultados de la crítica racional.

Esta es, brevemente esbozada, mi solución al problema de Hume, el problema lógico de la inducción[18]. Naturalmente, ha

[18] Para otras formulaciones, véase mi *L. I. C.*, *passim*, y especialmente mi carta al editor de *Erkenntnis*, reimpresa en el Apéndice *I de *L. I. C.* [Véase

habido otras formulaciones del problema, y hay aspectos de él que no han sido suficientemente analizados en esta sección. Por tanto, voy a proceder ahora a seguir el problema a través de algunos de sus aspectos y fases o etapas. El análisis de todo ello me llevará más allá del problema lógico de la inducción, tal como se ha formulado aquí, a lo que llamaré la cuarta fase o etapa, la etapa metafísica, del problema de la inducción.

Antes de tratar esos diferentes aspectos del problema de la inducción, necesitamos, primero, abordar algunas otras cuestiones que no son ni lógicas ni metodológicas ni metafísicas: tenemos que considerar algunos de los llamados «hechos» sobre la inducción y el aprendizaje.

3. SOBRE LOS LLAMADOS PROCEDIMIENTOS INDUCTIVOS, CON NOTAS SOBRE EL APRENDIZAJE Y SOBRE EL ESTILO INDUCTIVO

I

Parece que casi todo el mundo cree en la inducción; esto es, cree que aprendemos por la repetición de las observaciones. Incluso Hume, a pesar de su gran descubrimiento de que una ley natural no puede ni establecerse ni tampoco hacerse «probable» por inducción, continuó creyendo firmemente que los animales y los hombres aprendían por la repetición: por medio

---

formulaciones posteriores en *Conocimiento objetivo (Obj. Kn.)*, capítulos 1 y 2 *Ed.*] Puede mencionarse aquí una de las interpretaciones erróneas que ha encontrado mi teoría. (Forma la base de una reseña de G. J. Warnock de *L. I. C.* en *Mind*, 69, pp. 99-101.) En la introducción al capítulo X de *L. I. C.* escribí: «tendríamos que intentar la averiguación de hasta qué punto [una hipótesis] ha sido capaz de demostrar que es apta para sobrevivir —y ello por haber salido indemne de las contrastaciones—». Como lo muestra el contexto de toda la discusión, no quería dar a entender, naturalmente, que una teoría que ha sobrevivido *hasta ahora* haya probado *su aptitud para sobrevivir a contrastaciones futuras* (como supone la reseña). Por el contrario, dije una y otra vez, con énfasis, que si tenemos que elegir, elegimos, por el momento, la teoría que parece ser la mejor a la luz de la crítica, incluyendo las contrastaciones; y que *esta elección es perfectamente razonable, aunque no podemos saber si la teoría sobrevivirá contrastaciones futuras* (que pueden ser diferentes de las pasadas o no serlo) y aunque podemos temer —o esperar— que no sobrevivirá las contrastaciones. (Por cierto, el reseñista cae en un error de hecho cuando dice «Popper pone énfasis» en que el problema de la inducción es «insoluble»; porque ofrecer una solución al problema —incluso una que no sea una justificación de la inducción— no es lo mismo que decir, enfáticamente o de otro modo, que el problema no tiene solución. Este error está relacionado con el hecho de que lo que yo llamo «el problema de la inducción» o «el problema de Hume» es totalmente diferente de lo que el reseñista dice que yo he llamado así.)

de observaciones repetidas, así como por medio de la formación de hábitos, o el fortalecimiento de hábitos, por repetición. Y sostuvo la teoría de que, aunque la inducción era indefendible racionalmente y no daba como resultado nada mejor que una creencia no razonada, a pesar de ello, podía confiarse en ella en lo principal; que era, en cualquier caso, más digna de confianza y más útil que la razón y el proceso del razonamiento; y que la «experiencia» era, pues, el resultado no razonado de una acumulación (más o menos pasiva) de observaciones.

En contra de todo esto, yo pienso que en realidad *nunca* pueden hacerse inferencias inductivas o hacer uso de lo que ahora se llaman «procedimientos inductivos». Por el contrario, siempre descubrimos las regularidades por el método, esencialmente diferente, de prueba y error, de conjetura y refutación, o de aprender de nuestras equivocaciones; un método que hace el descubrimiento de regularidades mucho más interesante de lo que Hume pensó. El método de aprendizaje por prueba y error se ha tomado, equivocadamente, como un método de aprendizaje por repetición. La «experiencia» se gana aprendiendo de nuestros errores, más que por la acumulación o asociación de observaciones. Se gana mediante un enfoque activamente crítico: con el uso crítico de experimentos y observaciones concebidas para ayudarnos a descubrir en dónde nos hemos descaminado[19].

II

Por ello, aunque estoy de acuerdo con el análisis de Hume del problema lógico de la *validez* de la inducción —es decir, con su tesis de que la inducción no es válida— estoy en desacuerdo con él y, mucho me temo, con la mayoría de la gente, respecto a una cuestión puramente *de hecho*. Creo que *la alegación de que, de hecho, procedemos por inducción es puro mito*, y que la pretendida evidencia en favor de este pretendido hecho es un parte inexistente y en parte producto de una interpretación equivocada de los hechos.

Me apresuro a añadir, sin embargo, que mi tesis factual no atañe en absoluto a mis doctrinas lógicas o metodológicas o epistemológicas. Porque la cuestión factual, psicológica e histórica «¿cómo llegamos a nuestras teorías?», aunque puede ser fascinante[20], no concierne a la cuestión lógica, metodológica y

[19] Véase también, más adelante en esta sección, subsección IX, nota 27 (y texto).

[20] Jacques Hadamard ha escrito un interesantísimo libro sobre este tema: *The Psychology of Invention in the Mathematical Field*, 1945.

epistemológica de la validez. En esto vuelvo a seguir a Hume. En efecto, el mayor logro de Hume fue separar rigurosamente esos dos problemas. Al darles respuestas casi contrarias dejó perfectamente claro que eran totalmente distintos.

Algunos científicos encuentran, al parecer, que se les ocurren las mejores ideas mientras fuman; otros, mientras beben café o whisky. Por tanto, no hay razón para que yo no admita que a algunos puedan ocurrírseles mientras observan o repiten observaciones. Y, en este sentido, estoy bien dispuesto a mitigar mi tesis de que nunca procedemos por inducción: sustituyamos «nunca» por «casi nunca».

Pero, hecha esta concesión, quiero explicar que tanto si una teoría se nos ocurre por primera vez mientras fumamos, leemos y observamos o, incluso, mientras dormimos, la cuestión importante sigue siendo: ¿cuál es su valor lógico? ¿es una buena o mala teoría? Se ha dicho que a Darwin se le ocurrió su teoría de la selección natural cuando estaba leyendo a Malthus. Esta es una cuestión de considerable interés histórico, pero no tiene la más mínima relación con la cuestión del valor de la teoría de Darwin: incluso aunque la teoría de Malthus fuese verdadera y estuviese bien apoyada por la evidencia más fuerte, la de Darwin podría ser falsa y estar mal sustentada; e incluso aunque la teoría de Malthus fuese totalmente insostenible, la de Darwin podría ser excelente.

Naturalmente, todo el mundo admite esto. Y sin embargo, parece que quedan pocos filósofos que insistan en que tenemos que distinguir claramente entre cuestiones de *validez* (tales como si tenemos alguna *razón* para depender de la inducción) y cuestiones de *hecho* (tales como si realmente dependemos de la inducción o usamos «procedimientos inductivos»; o si una teoría se originó realmente por inducción, etc.).

III

La concepción más en boga sobre el asunto parece ser ésta: cuando Hume descubrió que la inducción no era válida, uso el término «no válida» [*invalid*] en el sentido de «no acorde con los cánones del razonamiento deductivo válido». Pero el descubrimiento de Hume es trivial, según la concepción sobre la que estoy informando, puesto que el razonamiento inductivo es una especie de razonamiento que es similar en algunos aspectos al razonamiento deductivo y diferente en otros. Por tanto, es bastante obvio que sus criterios o cánones no concordarán en *todo* aspecto con los del razonamiento deductivo. El razonamiento inductivo tiene sus propios criterios, sus propios

cánones, sus propios «procedimientos». Así, un «procedimiento inductivo» o una «inferencia inductiva» no será, en general, una «inferencia válida» —esto es, no será una inferencia *deductiva* válida—. Pero esto no es razón para que no sea «razonable», es decir, concorde con los criterios apropiados del razonamiento *inductivo:* es *inductivamente válida.*

Pero ¿qué son esos pretendidos criterios inductivos? Hay dos respuestas, dadas por dos escuelas filosóficas diferentes. Una escuela enseña que los criterios son los del razonamiento «probable», de acuerdo con las leyes del *cálculo de probabilidades.* He refutado esta teoría con todo detenimiento en mi *L. I. C.* y en otros lugares, y añadiré otras críticas más adelante en este *Post Scriptum.* La otra escuela enseña que debemos observar, y clasificar, los diversos «procedimientos inductivos» y los diversos usos de la palabra «probable» que pueden darse cuando hablamos de esos «procedimientos inductivos» y que los resultados de esas investigaciones nos permitirán establecer las leyes o criterios o cánones de la *validez inductiva.*

Así, según esta escuela, los «usos» factuales han de establecer los criterios. Ya no existe distinción entre hecho y criterio (entre las cuestiones de «*quid facti*» y «*quid juris*»)[21]. Ya no existe problema lógico de la validez o «justificación» de la inducción. No podemos «justificar» el uso de procedimientos inductivos, según esta escuela, de la misma manera que no podemos «justificar» el uso de procedimientos deductivos o de procedimientos culinarios o de cualquier otro procedimiento, excepto, quizá, en algún sentido prágmático de «justificar»; dicho de otro modo, por su éxito. Y los procedimientos inductivos no necesitan de una «justificación» más de lo que necesitan de ella los procedimientos deductivos o de cualquier otra índole.

Antes de pasar a criticar esta concepción de moda, quiero señalar que es, esencialmente, una concepción antigua, traducida a un lenguaje de moda. Porque la expresión antigua «las leyes del pensamiento», que solía referirse a los principios de la inferencia deductiva, contenía una alusión a la concepción de que no había modo de «justificar» la validez de la lógica deductiva excepto señalando *el hecho* psicológico de que, *de hecho,* pensamos de esta manera, o que, *de hecho,* estamos obligados a pensar de esta manera, u obligados a admitir que ciertas inferencias deductivas son inescapables o «necesarias»[22]. Una vez que la validez del razonamiento deductivo se redujo así a *hecho,* el camino estaba abierto para la aceptación de algún

[21] *Cf. L. I. C.,* sección 2, donde se hace referencia a Kant.
[22] *Cf. El desarrollo del conocimiento científico (C. & R.),* capítulo 9.

«principio de inducción» o de algunos «cánones del razonamiento inductivo» cuya validez era innecesario justificar de otro modo que por el hecho de que razonábamos, o quizá estábamos obligados a razonar, según esos cánones inductivos.

La versión actualmente de moda de esta antigua teoría sustituye un argumento psicológico totalmente incorrecto en favor de la identidad entre la cuestión de validez y la cuestión de hecho por un argumento pragmatista o conductista o de análisis del lenguaje, igualmente incorrecto.

Todos estos mecanismos son meros intentos de resucitar una teoría descartada hace largo tiempo por Hume y Kant; y los modernos resucitadores atropellan la obra de estos grandes hombres sin tratar de replicar a sus argumentos ni de entenderlos, siquiera. Es puro dogmatismo afirmar que no hay nada en la validez fuera del éxito. (Ninguna teoría científica tuvo más éxito que la de Newton. Si el éxito fuese nuestra preocupación, nadie la habría criticado, es decir, nadie habría reconsiderado la cuestión de su verdad. Sin embargo, la crítica que se le hizo llevó a una importante revolución intelectual; una revolución que tenía que haber revolucionado profundamente el pensamiento epistemológico.)

La identificación de modos válidos de pensamiento con el pensar real se ha intentado una y otra vez desde tiempo inmemorial. Uno de los más influyentes y perniciosos intentos en ese sentido fue la filosofía de Hegel de la identidad de la razón y la realidad. Las últimas teorías sobre la inducción equivalen a una renovación del intento de Hegel.

Pero los supuestos hechos que esas teorías de la inducción tratan de convertir en criterios, son enteramente imaginarios; y algunas personas cuyo conocimiento de la ciencia no es en absoluto despreciable —Albert Einstein, por ejemplo—[23] han negado que fuesen hechos. Pero incluso aunque todos nosotros, los que negamos la existencia de «procedimientos inductivos» estuviéramos equivocados, sería el súmmum del dogmatismo afirmar que esos discutidos «hechos» crean criterios de razonamiento cuya validez no está abierta a ulterior discusión. (Y aunque el arte de discutir críticamente sobre problemas filosóficos —y con él, la gran tradición del pensamiento filosófico racional— está desapareciendo rápidamente, yo, por mi parte, no estoy dispuesto a resignarme a este hecho sin tratar de cambiarlo; y estoy aún menos dispuesto a seguir la moda actual de hacer de este hecho un nuevo criterio de excelencia filosófica.)

[23] Véase la conferencia Herbert Spencer pronunciada por Einstein, *On the Metod of Theoretical Physics*, Oxford, 1933.

IV

No son los filósofos, naturalmente, los únicos que creen en la inducción y en la existencia de «procedimientos inductivos». Como ya dije antes, casi todo el mundo cree en ella, incluidos muchos psicólogos, sobre todo los interesados en la teoría del aprendizaje, muchos biólogos y unos cuantos físicos. A los físicos no les diré nada más aquí, puesto que ya me he referido a Einstein. Pero diré algo de la «teoría del aprendizaje» y también sobre lo que he llamado el «estilo inductivo», una manera de escribir que todavía sigue siendo el estilo obligado en muchas revistas de biología. (Véanse las subsecciones X y XI, más adelante.)

V

Respecto a la teoría del aprendizaje, *sin duda podemos aprender y aprendemos a partir de la experiencia.* Estoy incluso dispuesto a decir, con Hume y otros empiristas clásicos, que *todo aprendizaje es aprendizaje a partir de la experiencia.*

No obstante, cuando llegamos a la interpretación de la tesis «todo aprendizaje es aprendizaje a partir de la experiencia», difiero de Hume y de otros empiristas clásicos. Difiero radicalmente de ellos en la valoración del papel que juega la *repetición* en el proceso del *aprendizaje* y también en la valoración del papel que juega la *observación* en la adquisición de *experiencia.*

En cuanto al aprendizaje y la repetición, el no distinguir entre tres actividades enteramente diferentes, llamadas «aprendizaje» las tres, ha sido fuente de serias confusiones. (No quiero decir que no puedan distinguirse, con provecho, más de tres.) Las llamaré 1) aprendizaje por prueba y error (o por conjetura y refutación), 2) aprendizaje por medio de la formación de hábitos (o aprendizaje por repetición propiamente dicha) y 3) aprendizaje por imitación (o por asimilación de una tradición). Los tres tipos se encuentran tanto en los animales como en los humanos y juegan sus diversos papeles característicos en la adquisición de habilidades, así como de conocimiento teórico tal como el aprendizaje sobre nuevos hechos.

1) Sólo la primera de estas tres formas de aprendizaje, *aprendizaje por prueba y error* o por conjetura y refutación, atañe al aumento de nuestro conocimiento; sólo él es «aprendizaje» en el sentido de adquisición de *nueva* información: de descubrimiento de *nuevos* hechos y *nuevos* problemas, prácticos

y teóricos, y *nuevas* soluciones a nuestros problemas, viejos y nuevos. Este tipo de aprendizaje incluye el descubrimiento de nuevas habilidades y de nuevas formas de hacer las cosas. En los procesos de aprendizaje en este sentido, la repetición mecánica (como la de la gota que agujerea la piedra) no tiene ningún papel. No es la impresión repetida en nuestros sentidos los que lleva a un nuevo descubrimiento, sino algo enteramente diferente: nuestros repetidos y variados intentos de resolver un problema que, si está sin resolver, continúa irritándonos. Aquí es esencial que esos «repetidos» intentos difieran entre sí y que repitamos el mismo intento sólo cuando nos parezca ser acertado y sólo para volver a ponerlo a prueba; es decir, para contrastar, si es posible, bajo diferentes condiciones, la hipótesis de que dicho intento lleva invariablemente a una solución acertada de nuestro irritante problema.

Aprender por prueba y error incluye aprender a partir de la observación sistemática, así como de la observación fortuita, aunque de formas diferentes. *La observación sistemática* parte siempre de un problema que tratamos de resolver, o de conjeturas que tratamos de contrastar: esto es lo que la hace sistemática. Incluso cuando tratamos de determinar cierto parámetro por medio de una medición sistemática, hay una hipótesis subyacente: la hipótesis de que existe un parámetro que es invariable respecto a ciertos cambios en las condiciones de nuestras mediciones. Sin alguna hipótesis así, sea conscientemente propuesta o inconscientemente supuesta, la observación no puede ser sistemática. Sin embargo, incluso una llamada «*observación fortuita*», a pesar de ser la forma menos ingeniosa de hacer descubrimientos, sigue siendo un caso del método de prueba y error. Prácticamente por cada ejemplo de «observación fortuita» hay un ejemplo de *refutación* de alguna conjetura o supuesto o expectativa, consciente o inconscientemente sostenida. Una «observación fortuita» es como una piedra inesperada en nuestro camino: tropezamos con ella sólo porque no la esperábamos o, más exactamente, por *lo que* esperábamos, aunque inconscientemente, era que el camino fuera liso. Así, las llamadas «observaciones fortuitas» o «descubrimientos accidentales» —es decir, los descubrimientos por tropezón— no son tan accidentales como puede parecer a primera vista.

Además, la mayor parte de los ejemplos usuales de descubrimientos fortuitos están basados en «interpretaciones inductivistas» erróneas: incluso los ejemplos auténticos de descubrimiento por tropezón parecen ser raros. Quizá la primera inmunización de Pasteur a pollos, contra el cólera de los pollos (1880), o el descubrimiento de la acción catalítica de las sales de mercurio en la conversión de la naftalina en ácido ftálico

(calentándolo con ácido sulfúrico)[24], se aproximen a ejemplos auténticos. Pero muchos otros no lo son.

Oersted, por ejemplo, estaba *buscando* desesperadamente algunas interacciones electro-magnéticas. Y Roentgen, cuando se le interrogó sobre su descubrimiento de los rayos X, explicó: «Estaba *buscando* rayos invisibles»[25], rayos que esperaba detectar (como en el caso de los rayos infrarrojos y ultravioleta) por medio de una pantalla fluorescente. (Esa es la razón de que la pantalla estuviera allí.) Hay que admitir que placas fotográficas no expuestas colocadas cerca de tubos de Crookes habían mostrado muchas veces, contra toda expectativa, señales de exposición, y esto *podría* haber llevado fácilmente a un auténtico «descubrimiento por tropezón» de los rayos X, pero todos los implicados, incluido el propio Crookes, no supieron calibrar la significación de esos signos de exposición. En cuanto a la radiactividad, Becquerel estaba buscando conscientemente nuevos rayos; y en su trabajo con sales de uranio, en particular, iba guiado por una hipótesis (equivocada), debida a Henri Poincaré. El descubrimiento de la penicilina no fue tampoco un descubrimiento fortuito porque el tipo de efecto bacterizada (o bacteriostático) observado por Fleming era bien conocido de él y de otros y, por tanto, no era ni siquiera «inesperado». (Véase la subsección X, más adelante.) Además Fleming era muy consciente, incluso antes de su descubrimiento, de la posible significación de este tipo de efecto para fines terapéuticos.

Con los animales, como con los hombres, aprender originalmente por prueba y error resulta, no tanto en nuevos «conocimientos» como en nuevas habilidades. Sin embargo, las habilidades están ligadas a *expectativas*, conscientes o inconscientes: y el elemento de *error*, dentro del método de prueba y error, se manifiesta siempre en la frustración de alguna espectativa. (Lo que llamamos nuestro conocimiento —«conocimiento» en sentido subjetivo— puede decirse que consiste en nuestras expectativas conscientes, quizá incluso formuladas ver-

[24] Parece ser que esta historia de descubrimiento, junto con algunas variantes de ella, se ha convertido en una especie de leyenda entre los químicos. Según una comunicación privada del Profesor Alexander Findlay, el descubrimiento lo hizo, en 1896, Sapper, un joven químico alemán que rompió accidentalmente un termómetro en la mezcla y percibió una gran aceleración en la velocidad de la reacción. Es importante añadir que esta aceleración contenía la solución a un problema urgente de la química industrial que él estaba tratando de resolver. De modo que hubo un accidente: pero estuvo precedido y preparado por el problema del que era solución, de forma que la observación estuvo lejos de ser accidental. (Véase también A. Fleming, *Chemistry in the Service of Man*, 7.ª edición, 1947, pp. 318 ss.)

[25] Véase Otto Glasser, *Wilhelm Conrad Roentgen and the Early History of the Roentgen Rays*, 1933, p. 13.

balmente.) Como ejemplos de nuevas habilidades adquiridas por prueba y error podemos citar tocar el piano o montar en bicicleta; o, más exactamente, encontrar la mejor manera de mover los dedos en un pasaje de piano, o aprender cómo evitar una caída cuando se está montando en bicicleta. En esos casos, primero tratamos de resolver un problema conscientemente por medio de pruebas sistemáticas —por medio de la eliminación o rechazo (o falsación) de las soluciones insatisfactorias—; las repeticiones posteriores cumplen una función diferente. Los casos son, por tanto, útiles para comparar el carácter y la función de la repetición en los tipos de aprendizaje diferenciados aquí como 1) y 2).

2) El segundo tipo de aprendizaje —*formación de hábitos por medio de la repetición propiamente dicha* (o por medio de la repetición «mecánica»)— debe distinguirse claramente del primero. Aquí no buscamos una nueva solución a un problema, sino que tratamos de familiarizarnos con una solución descubierta previamente por el método de prueba y error (o aprendida por imitación; véase más abajo, en 3)).

Para mostrar cómo se diferencia este tipo de aprendizaje del primero, es instructivo considerar nuestros principales ejemplos, el de aprender a tocar un cierto pasaje al piano, y el de aprender a no caerse de la bicicleta.

Hay pocas destrezas humanas en las que la «práctica» constante —es decir, no sólo la repetición, sino la repetición más o menos «mecánica»— sea tan importante como en el aprendizaje del piano. Sin embargo, no descubrimos nada nuevo practicando: por ejemplo, no aprendemos un nuevo movimiento de los dedos. Sólo después de haber descubierto un nuevo movimiento de los dedos por prueba y error, esto es, después de compararlo con soluciones alternativas del problema y rechazar las soluciones menos convenientes, sólo entonces podemos empezar a «practicarlo». Así, la función de la repetición mecánica —de «practicar» o de «aprender de memoria»— no es descubrir nada nuevo, sino familiarizarse con algo descubierto previamente. Su función no es hacernos conscientes de un nuevo problema (como es la función de contrastar repetidamente algunas soluciones provisionales), sino eliminar, en la medida de lo posible, el elemento de conciencia de nuestra ejecución. Así, alcanzamos un estado en el que el problema original —por ejemplo el de coordinar la partitura y el movimiento de los dedos— se desvanece por completo y podemos entregar toda nuestra atención a algo más importante: a la idea musical, al fraseo del pasaje. La función de la «práctica» o de la repetición propiamente dicha, en el aprendizaje de la bicicleta (o en el de la conducción de un coche) es la misma: no produce des-

cubrimiento, ni siquiera una *nueva* habilidad, aunque puede transformar un descubrimiento (un descubrimiento de cómo hacer las cosas en una nueva *habilidad*), y al hacer inconscientes algunas acciones, nos deja libres para prestar nuestra atención a los problemas de la circulación.

La doctrina inductivista de que todo aprendizaje es aprendizaje de memoria y que incluso el aumento de nuestro conocimiento es el resultado de la formación de hábitos por medio de la repetición mecánica, es, por tanto, equivocada. La repetición en cuanto tal no puede atraer nuestra atención; por el contrario, tiende a hacer inconscientes nuestras expectativas. (Podemos no oír el tic-tac del reloj, pero podemos «oírlo» pararse.)

La popularidad de la idea de que nuestro conocimiento aumenta a través de la inducción por repetición se debe sin duda a una confusión de los dos sentidos, 1) y 2), del aprendizaje.

3) Pero mencionaré brevemente también una tercera clase de aprendizaje: aprendizaje por imitación. Es una de las formas de aprendizaje más primitivas y más importantes; y en ella, la enormemente compleja base instintiva del aprendizaje y la función que juegan la sugestión y las emociones son más obvias que en otras formas de aprendizaje (aunque éstas están, naturalmente, siempre presentes). Lo importante para nuestra discusión es que, desde el punto de vista del aprendiz individual, aprender por imitación es siempre un proceso típico de prueba y error: un niño (o un animal joven) *prueba*, consciente o inconscientemente, a imitar a su padre y, o se corrige a sí mismo, o es corregido por su padre. Este proceso de prueba y error constituye la etapa primera y fundamental del proceso de imitación. Es, pues, una etapa de *descubrimiento:* el niño *descubre* por imitación cómo se anda; y esto quiere decir, en parte, por el método de prueba y error. Esta etapa puede ir seguida, naturalmente, por una en la que la habilidad recién descubierta, como consecuencia de «practicarla», se ejecute inconscientemente y se convierta así en un hábito.

VI

El inductivista interpreta el conocimiento («conocimiento» en el sentido subjetivo) como consistente en expectativas; y yo también. Pero él va más allá, interpreta una expectativa como el recuerdo de observaciones ligadas por asociaciones que son el resultado de la repetición. (El piensa que el perro espera comida cuando suena el timbre solamente porque en repetidas

ocasiones anteriores la comida llegó después de que sonase el timbre.)

Contrariamente a esto, yo creo que con la prueba y el error se crean nuevas expectativas: nos creamos expectativas provisionales en campos de interés, en campos en los que tenemos problemas, en campos en los que somos capaces de aprender, es decir, de corregir nuestras expectativas. Si las expectativas recién formuladas tienen éxito, pueden convertirse (por repetición) en automáticas, inconscientes y petrificadas, y gradualmente dejamos de ser capaces de aprender en ese campo determinado. La obtención de alimento es un terreno en el que el comportamiento del perro es normalmente «plástico»: puede aprender en este terreno y descubre regularidades por el método de prueba y error, por anticipación y refutación. De este modo también se crea al principio la teoría de que la comida aparecerá cuando suene el timbre; después la anticipación o expectación puede convertirse en habitual y petrificarse con la repetición.

VII

La concepción aquí presentada es totalmente opuesta a las teorías de la asociación y del reflejo condicionado (véase la sección 9, más adelante). La primera de ellas funciona suponiendo que se nos «dan» términos de asociación simples o «datos» simples. Esta suposición es ingenua e insostenible, incluso aunque dejemos de lado la situación de estímulo complejo y supongamos que los estímulos son simples. La segunda de esas teorías es muy diferente y menos objetable en la medida en que no hace suposiciones sobre datos o términos simples, sino que trata de explicar cómo unos estímulos (que pueden ser sumamente complejos) originalmente sin relación con una expectativa (que puede ser también sumamente compleja) pueden convertirse en señales capaces de desatar precisamente esa expectativa. Esto es bastante aceptable como problema; pero la solución, la teoría de un reflejo condicionado, es inaceptable. Supone la existencia de reflejos elementales, no complejos, no condicionados, a partir de los cuales se crea el reflejo condicionado y supone que hay que explicar todo aprendizaje como el condicionamiento de reflejos. Ambas suposiciones son erróneas. El aprendizaje en el sentido primero y fundamental, es decir, el aprendizaje por prueba y error, contiene un elemento de invención, de acción creativa, que va mucho más allá de cualquier simple reflejo; y hablar aquí, donde la invención y la plasticidad en la acción son de la mayor importancia, del «condicionamiento de los reflejos» sugiere equivocadamente que todo

aprendizaje, incluidos el descubrimiento y la invención, puede explicarse por la repetición. Pero hemos visto que el aprendizaje por repetición es menos fundamental (y, por tanto, también menos elemental) que el aprendizaje por prueba y error[26].

VIII

El error del inductivista no se limita a su falta de apreciación de la diferencia entre aprendizaje por prueba y error y aprendizaje de memoria, o a su suposición consiguiente de que podemos aumentar nuestro conocimiento con la formación de hábitos. También cree que existe una materia prima para el conocimiento en forma de percepciones u observaciones o impresiones sensibles o «datos» sensibles que nos son «dados» desde el mundo exterior sin que nosotros intervengamos. Esta es una teoría psicológica insostenible, ampliamente refutada por los hechos. En el cine, lo que nos es «dado» es una sucesión de escenas inmóviles, pero lo que vemos y observamos o percibimos es movimiento; y no podemos evitar ver el movimiento, incluso aunque *sepamos* que no estamos viendo más que fotografías de (digamos) un dibujo animado.

El simple hecho es que ver o percibir u observar es una reacción, no solamente ante los estímulos visuales, sino ante ciertas situaciones complejas en las que no sólo juegan un papel los complejos y las sucesiones de estímulos, sino también nuestros problemas, temores, esperanzas, nuestras necesidades y satisfacciones, nuestros gustos y aversiones. Nuestra reacción —esto es, nuestra experiencia perceptual inmediata— es influenciada por todo esto y también, en gran medida, por nuestro conocimiento previo, por nuestras expectativas o esperanzas, que proporcionan una especie de marco esquemático para nuestras reacciones. Si aprendemos, en el sentido de añadir a nuestro conocimiento, por medio de nuestras observaciones o percepciones, entonces lo hacemos porque observar o percibir consiste en modificar, precisar, corregir y, con frecuencia, falsar nuestras esperanzas. Así, la teoría inductivista es siempre superficial: un análisis más riguroso muestra que lo que el inductivismo toma ingenuamente como un «dato» de nuestros sentidos

[26] [*He publicado varias críticas del llamado «reflejo condicionado». Brevemente: he afirmado que los perros de Pavlov no estaban «condicionados», sino que se formaron una teoría —en un terreno (obtención de alimento) en el que la formación de teorías es de una importancia vital— según la cual la comida llega cuando suena la campana. Véase especialmente mi contribución a Karl R. Popper y John C. Eccles, *El yo y su cerebro (S. & I. B.*, 1977) pp. 91, 135-38; y los artículos de Roger James a los que se hace referencia ahí y, ahora, también el excelente libro de James, *Return to Reason*, Somerset, 1980.]

consiste, en realidad, en un complejo intercambio entre el organismo y su entorno: el proceso de modificar o corregir nuestras esperanzas y de refutar nuestras conjeturas, que es tan característico de todas las formas de aprendizaje con las que añadimos a nuestro conocimiento[27].

IX

La concepción empírica clásica —que la experiencia es consecuencia de aprender por medio de la repetición de la observación— es un sistema cerrado de prejuicios cuyo análisis crítico es resistido y con frecuencia resentido por sus defensores. Este sistema de prejuicios es muy popular y ha pasado a formar parte del «sentido común»; sin embargo, también puede ser descrito como un *sistema filosófico* docto y algo artificial porque hay muchas indicaciones de que «experiencia» y «aprender por experiencia» se usan corriente y popularmente en el sentido del método de prueba y error: «aprender por experiencia» significa, bastante corrientemente, «aprender de nuestros errores» (en lugar de «aprender de memoria» o «aprender por la asociación de observaciones», por no hablar de «datos sensibles»). Como lo expresa Oscar Wilde: «Experiencia es el nombre que todos dan a sus errores»[28].

Brevemente, lo que es realmente el uso ordinario o común o popular de la palabra «experiencia» entraña una teoría que es similar a la mía, expuesta con considerable extensión en *L. I. C.* y en otros lugares[29]. Según esta teoría, la experiencia no debe tomarse como la «fuente» última «del conocimiento», sino, más bien, como un sistema de expectativas o de esperanzas falibles a las que cada uno llega por prueba y error. Sin embargo, mi referencia a los usos populares —a la etimología— de «experiencia» no intenta ser un argumento en favor de mi propia concepción, porque tales usos inducen a error con frecuencia (aunque no debemos apartarnos del sentido común

[27] Compárese con todo esto el capítulo 1 de *El desarrollo del conocimiento científico (C. & R.)* [*Cf. Conocimiento objetivo (Obj. Kn.)*, capítulo 7. *Ed.*]

[28] Oscar Wilde, *El abanico de Lady Windermere*, Acto III. *Cf.* sección 15 del Addendum (1961) al tomo II de *La sociedad abierta (O. S.*, 4.ª edición), pp. 388 y ss. Véase también, por ejemplo, el artículo «Experience» en el *Oxford English Dictionary*, que, naturalmente, también incluye el significado filosófico o epistemológico, «la observación de hechos o sucesos considerada como fuente de conocimiento» (citando a Thomas Reid), pero que, de otro modo, sostiene la concepción de que (aparte de la «experiencia religiosa») «experiencia» significa, por lo común, el resultado de aprender de nuestros errores.

[29] Aparte de la sección a la que se refiere la nota anterior, véase también la Introducción a *El desarrollo del conocimiento científico (C. & R.)* y *Conocimiento objetivo (Obj. Kn.)*, capítulo 2.

—sea lo que fuere lo que eso significa— sin una razón suficientemente buena). Mi intención es, más bien, señalar que los analistas de los usos lingüísticos ordinarios no se mantienen firmes en sus opiniones cuando, llevados por su filosofía inductivista de la experiencia como fuente de conocimiento, dan por sentada la existencia de «procedimientos inductivos».

X

El inductivismo de muchos biólogos es, creo, tradicional, y data de Bacon, Boyle, Leeuwenthoek, y de los primeros tiempos de la Royal Society. Traté de esbozar la filosofía que hay detrás de esta tradición —la creencia de que la naturaleza es un libro abierto que debe leerse sin prejuicios— en la última sección de mi *L. I. C.* y, con mucho más detalle, en la introducción de *El desarrollo del conocimiento científico: Conjeturas y refutaciones.*

En ningún lugar es tan conspicuo el poder de la tradición inductivista como en lo que he llamado «*el estilo inductivo*», una cierta forma de presentar las investigaciones propias que es todavía la manera tradicional de escribir en una serie de revistas biológicas, aunque a estas alturas ya ha desaparecido casi por completo de todas las revistas de física o de química. La idea básica que inspira el estilo inductivo es ésta: debemos atenernos cuidadosamente a nuestras observaciones reales y debemos abstenernos de teorizar, porque eso puede hacernos adquirir prejuicios teóricos que pueden sesgar o viciar nuestras observaciones fácilmente, si no tenemos *mucho* cuidado.

Por esta razón, un artículo escrito en estilo inductivo tiene, esencialmente, la siguiente estructura:

1) En primer lugar explica las preparaciones para nuestra observación. A ellas pertenecen, por ejemplo, las disposiciones experimentales, tales como el aparato usado, su preparación para el experimento y la preparación de los objetos de la observación.

2) La parte principal del artículo consiste en una descripción pura, imparcial desde el punto de vista teórico, de los resultados del experimento: las observaciones realizadas, incluidas las mediciones (si las hay).

3) Sigue un informe de las repeticiones del experimento con una valoración de la fiabilidad de los resultados o los errores probables. (En los últimos tiempos, esto puede incluir algunas estadística.)

4) Optativo: una comparación de los resultados con otros anteriores o con los de otros que trabajan en el mismo campo.

5) También optativo: sugerencias para futuras observacio-

nes, para mejoras deseables en el aparato y para nuevas mediciones.

6) El artículo se concluye (una vez más, esto es optativo) con un breve epílogo, a menudo unas cuantas líneas nada más, y a veces en letra pequeña, conteniendo la formulación de una hipótesis sugerida por los resultados experimentales del artículo.

No sugiero, naturalmente, que los artículos se ajusten siempre rígidamente a estos puntos. Algunos pueden faltar y añadirse otros. Lo que sugiero es que hay una tendencia a hacer creer a los jóvenes biólogos que ésta es la forma adecuada de presentar los resultados y que incluso los maestros se ajustan a este tipo de presentación.

Sin duda, la idea que inspira el estilo inductivo —la idea de ajustarse estrictamente a los hechos observados y de excluir predilecciones y prejuicios— es loable. Y, sin duda, los que han sido acostumbrados a escribir con este estilo no son conscientes de que esta idea tan loable y aparentemente segura es, ella misma, el erróneo resultado de un prejuicio —peor aún, de un prejuicio filosófico— y de una teoría errónea sobre la objetividad. (La objetividad no es el resultado de la observación desinteresada y sin prejuicios. La objetividad, y también la observación imparcial, son el resultado de la crítica, incluida la crítica de las descripciones de observaciones. Porque no podemos evitar ni suprimir nuestras teorías, ni impedir que influencien nuestras observaciones; pero podemos intentar reconocerlas como hipótesis y formularlas explícitamente, de modo que puedan ser criticadas.)

Como ejemplo del estilo inductivo puedo mencionar aquí el artículo clásico en el que Alexander Fleming informó del descubrimiento de la penicilina[30]. En él describe su observación de la invasión accidental de un cultivo de bacterias por un agente que las destruye. Es una descripción de observaciones, de lo que ocurrió; y aunque no *dice*, naturalmente, que este tipo de suceso fuera inesperado, ni que ocurriese por primera vez, su estilo inductivo suele dejar al inocente lector (por ejemplo, un filósofo) con la impresión de que no sólo la invasión del cultivo de bacterias por la penicilina fue accidental (esto lo fue, en un sentido), sino que además no se esperaba.

Pero los hechos históricos muestran que esta impresión sería equivocada. Por lo menos desde Metchnikov (1845-1916), las teorías sobre «anticuerpos» —es decir, células o moléculas u otros agentes microscópicos que comiesen o destruyesen o inhibiesen el crecimiento de bacterias peligrosas— se habían discu-

[30] A. Fleming, *British Journal of Experimental Pathology*, 1929, pp. 226 ss.

tido constantemente entre los bacteriólogos. Tampoco fue el moho bactericida de Alexander Fleming el primero al que se había observado establecerse accidentalmente en un cultivo de microbios. De hecho, los bacteriólogos habían sostenido durante mucho tiempo la esperanza de que, de este modo, podrían llegar algún día a descubrir un método eficaz para matar a las bacterias en el hombre. En 1924, Sinclair Lewis había publicado *Arrowsmith*, una novela muy buena en la que un incidente muy similar al descrito en el artículo de Fleming tiene un papel importante. (Las partes sobre bacteriología fueron escritas en colaboración con el Dr. Paul de Kruif, que más tarde se hizo famoso con su *Microbe Hunters*, una historia popular y muy legible de la bacteriología que contiene, además, un análisis muy bueno de los métodos del descubrimiento científico.)

En realidad, en aquel momento se conocían muchos incidentes similares, y muchas sustancias que eran similares *prima facie* a la penicilina; el principal problema era si alguna de ellas sería apropiada para fines médicos. Fleming conjeturó que la penicilina lo sería. Pero tardó una década en conseguir la colaboración, muy necesaria, de una químico competente. Una década después del descubrimiento de Fleming, Howard Florey y sus colaboradores descubrieron los sorprendentes poderes curativos de la penicilina, confirmando así la conjetura de Fleming. Pero incluso esos sorprendentes poderes no eran totalmente inesperados, ya que Paul Ehrlich (1854-1915) había tenido la esperanza de encontrar una sustancia así de poderosa y las medicinas a base de azufre (cuya acción parece ser algo similar a la de la penicilina) fueron inventadas por investigadores formados en la tradición de Ehrlich.

Así pues, el descubrimiento de Fleming no fue accidental, realmente: fue la obra de un gran descubridor que sabía muy bien lo que estaba haciendo y lo que merecía la pena describir: y aunque fue un accidente que el hongo cuyas propiedades antibióticas había observado resultase no ser tóxico, se esperaba y se deseaba que existieran sustancias de este tipo desde hacía mucho tiempo. Esta esperanza motivaba tanto el trabajo de Fleming como el del equipo de Florey.

Sin embargo, buscamos en vano, en los primeros artículos, un enunciado de esos motivos, esperanzas y expectativas; de los problemas que hacen significativos los artículos. Debido al estilo inductivo de sus publicaciones, las esperanzas y expectaciones especulativas de los biólogos se transmiten normalmente por medio de una tradición oral en lugar de por escrito. En cualquier caso, pocas revistas biológicas estarían dispuestas a aceptar un artículo que expusiera ese tipo de especulaciones

teóricas, porque violan las normas aceptadas del estilo inductivo[31].

XI

La crítica del estilo inductivo no estaría completa si faltase la sugerencia de algo con que sustituirlo. Un científico debe ser libre, naturalmente, de presentar sus resultados como le parezca oportuno. Pero un estudioso de la lógica de la ciencia debe ser libre asimismo de someter a crítica un tipo de estructura general que los científicos pueden hacer bien en adoptar a menos que se les den buenas razones para apartarse de ella. Un trabajo experimental típico debería estar organizado, según mi plan, como sigue:

1) Una exposición clara del problema o, si puede suponerse que el problema es bien conocido, una referencia clara a él y a una exposición de él. El autor debería dejar claro también si acepta la situación del problema tal como la esboza algún predecesor o si ve el problema de forma diferente. Esto daría al autor la oportunidad de clarificar para sí mismo (y quizá para otros) la siempre cambiante *situación del problema.*

2) Un examen más detallado de las hipótesis pertinentes que conciernen al problema (y a los experimentos relacionados con las hipótesis, indicando el grado en que dichos experimentos pueden contribuir a la evaluación de las hipótesis).

3) Un enunciado más específico de la (o las) hipótesis que el autor intenta proponer o discutir o contrastar experimentalmente.

4) Una descripción de los experimentos y sus resultados.

5) Una evaluación: si la situación del problema ha cambiado; y si es así, cómo.

6) Sugerencias para trabajos posteriores que surjan del trabajo descrito.

Estos puntos parecen bastante obvios, si se adopta mi concepción sobre el procedimiento de la ciencia. No obstante, pocos autores adoptan nada parecido a este «estilo»: la tradición inductivista es demasiado fuerte.

Sin embargo, cada día es más urgente que todo artículo explique su localización dentro de la situación del problema en las diversas ciencias. El asombroso aumento de publicaciones junto a la disminución de los contactos personales entre los que

[31] [Véase Popper, «Science: Problems, Aims, Responsibilities», *Federation Proceedings* 22, julio-agosto 1963, pp. 961-972, esp. pp. 970-71. Este artículo, así como el de Peter Medawar, «Is he Scientific Paper a Fraud?» (*BBC* 3.er Programa, 1963-64), han llevado a un cambio. *Ed.*]

trabajan en campos *vecinos*, ha llevado a una atomización de la ciencia que hace que sea más importante que nunca resaltar la importancia de los problemas teóricos y de las teorías. Porque las teorías constituyen la red de coordenadas de la ciencia.

XII

He mencionado algunas razones que, aunque no es probable que convenzan a un inductivista de la inexistencia de procedimientos inductivos, al menos deberían mostrar que hay interpretaciones alternativas de los hechos, interpretaciones que él descuida con peligro para sí mismo; que la creencia en la existencia de procedimientos inductivos puede ponerse en tela de juicio; y que incluso aunque hubiera algunos ejemplos de procedimientos inductivos *prima facie*, un análisis más profundo puede mostrar que son ejemplos de un método cuya estructura lógica es enteramente diferente: el método de prueba y error.

Así, ni el recurso al supuesto *hecho* de que usamos la inducción o procedimientos inductivos, ni los intentos de convertir este *hecho* en un criterio, pueden aceptarse en lugar de argumentos: el inductivista que trata de justificar nuestras teorías o creencias mediante el recurso o la inducción no puede evitar la tarea de dar una justificación racional de la inducción.

XIII

El método de inducción por repetición trata de proporcionar un criterio de justificación. (Por otro lado, el método de prueba y error —de aprender con nuestras equivocaciones— es puramente crítico, selectivo. Sólo se convierte en justificacionista o inductivista si suponemos, equivocadamente, con Bacon o Mill, que es posible justificar una teoría mediante la eliminación completa de todas sus alternativas; pero el número de alternativas incontrastadas es siempre infinito[32], y siempre hay posibilidades en las que no se ha pensado.) El inductivista es, esencialmente, un justificacionista (en el sentido de la sección 2.)

Hume pensó que la inducción era *racionalmente* injustificable, pero que tenía su propio tipo de justificación; se justifica a sí misma en la práctica a causa de su alto grado de fiabilidad, en el que no podemos dejar de creer, aunque sólo de forma

[32] Esto lo enunciaron claramente Jeffreys y Wrinch, *Phil. Mag.* 42, 1921, pp. 369 ss.; véase *L. I. C.*, p. 132, nota *1; véase también *Die beiden bleme*, esp. Grundpropp. XIX s., nota 11.

irracional. Los inductivistas modernos también piensan que la inducción, aunque obviamente distinta de la deducción, tiene su propio tipo de justificación: se justifica a sí mismo en la práctica y establece sus propios patrones: se auto-confiere validez y autenticidad. En la inducción, los hechos y los criterios son conciliables, como lo son en Dios. La diferencia lógica entre Hume y la concepción moderna es, pues, la diferencia entre un creyente escéptico y algo recalcitrante y un teísta dogmático.

Pero la racionalidad —es decir, la crítica— y el dualismo entre hechos y criterios puede salvaguardarse fácilmente si renunciamos al justificacionismo. Si lo hacemos, nos hacemos también conscientes de la brecha lógica entre la inducción por repetición y el método de prueba y error o de aprender de nuestras equivocaciones.

Los creyentes en los procedimientos inductivos tienen una cosa en común conmigo: los dos diferimos de Hume en que pensamos que, de algún modo, los hechos o procedimientos para adquirir conocimiento están estrechamente relacionados con algunos criterios de racionalidad. Pero mientras ellos creen que los hechos, o lo que ellos piensan que son los hechos, se auto-confieren autenticidad y crean sus propios criterios de racionalidad inductiva, yo conjeturo que, de manera general, los criterios racionales (que yo creo que son los criterios de la crítica, es decir, los criterios lógicos) probablemente determinan nuestros procedimientos. Y como la razón y la lógica nos dicen que, racionalmente, no hay inducción ni hay justificación, sino sólo crítica y eliminación, es una buena idea ver si esos hechos del descubrimiento científico no pueden interpretarse —e interpretarse mejor— como procedimientos de prueba y error.

4. UNA FAMILIA DE CUATRO PROBLEMAS DE INDUCCION

La formulación del problema de Hume de la inducción dada en la sección 2 es, creo yo, de importancia fundamental. Pero hay otras formulaciones del problema que resaltan aspectos diferentes de él y que pueden considerarse como cuatro fases o etapas diferentes de su discusión.

En esta sección me propongo distinguir cuatro fase de ese tipo:

1) Una ligera variante de la formulación dada en la sección 2. Puede llamarse el desafío de Russell y puede formularse como la pregunta: *¿cuál es la diferencia entre el lunático y el científico?* Está estrechamente relacionada con el «problema de demarcación», es decir, el problema de encontrar una caracteri-

zación adecuada de la naturaleza empírica de las teorías científicas.

2) El llamado «*problema de la creencia racional*».

3) La cuestión de si podemos hacer inferencias sobre el futuro o si el futuro será como el pasado; una cuestión que el propio Hume no distinguió suficientemente del problema de la inducción. Lo llamaré «*el problema de Hume sobre el mañana*». Estas tres cuestiones tienen, todas ellas, un carácter lógico, o epistemológico o metodológico y trataré de mostrar en esta sección que no hacen falta ideas nuevas para darles solución.

Pero nuestra exposición nos llevará a distinguir una cuarta fase o etapa que, a pesar de una aparente similaridad con la tercera, es muy diferente de ella en su carácter lógico. Puede llamarse:

4) «La fase metafísica del problema del mañana» o «*la fase cuarta o metafísica del problema de la inducción*». La discusión de esta cuarta fase del problema se abordará en la sección 5. En esta sección me limitaré a una discusión de las tres primeras fases o etapas.

I

Bertrand Russell fue el primer filósofo, desde Kant, que sintió el problema de Hume de la inducción en toda su fuerza. Kant había visto claramente que, si Hume tenía razón, el conocimiento de carácter universal y, por tanto, pensaba él, el conocimiento científico, no podía existir; pero como creía que el ejemplo de las matemáticas y, más importante aún, el de la mecánica de Newton, mostraban que poseíamos en efecto cierto conocimiento científico, Kant pensaba que el problema central de la filosofía era explicar cómo era posible que existiera; es decir, explicar por qué estaba equivocado Hume.

Russell comprendió el problema del mismo modo, aunque la detallada solución que le dio difería considerablemente de la de Kant. (Por ejemplo, al describir las leyes de la mecánica como mero conocimiento probable, en oposición al conocimiento cierto de las matemáticas, Russell ensanchó aún más la brecha existente entre la ciencia matemática y la física, una brecha de la que ya había tratado Kant.)

Russell discute largamente la inducción en muchos lugares; primero, creo, en su incomparable libro *The Problems of Philosophy* (1912). Aunque en este delgado pero gran libro no se refiere al Hume como el que originó el problema, lo hace en su *History of Western Philosophy* (1946). Allí, en su capítulo sobre Hume, Russell formula el problema como sigue:

Si Hume tiene razón en que *no podemos hacer ninguna inferencia válida de la observación a la teoría,* entonces nuestra creencia en la ciencia no es ya razonable. Porque cualquier supuesta teoría científica, por arbitraria que sea, se convierte en tan buena —o tan justificable— como cualquiera otra, porque *ninguna* es justificable: la frase «mi suposición vale tanto como la suya» gobernará el método científico como principio único. Así, si Hume tuviera razón, «*no habría diferencia entre la locura y la cordura*»[33], y las obsesiones y delusiones de los locos serían tan razonables como las teorías y los descubrimientos de un gran científico.

Una sencilla y virtualmente completa respuesta a este desafío de Russell está implícita en las discusiones de la sección 2. Hay que admitir que no podemos justificar la afirmación de que la teoría del científico es verdadera; igual, de hecho, que la afirmación de que las delusiones no son tales. Pero podemos defender la afirmación de que la teoría del científico es mejor: mejor incluso en el sentido algo estrecho de que está mejor sustentada por las observaciones. Porque las observaciones *pueden* ser cruciales entre dos teorías, en el sentido de que pueden contradecir a una y ser compatibles con la otra. El argumento de Hume no establece que *no podamos hacer inferencias de la observación a la teoría:* establece solamente que no podemos hacer inferencias *verificadoras* de las observaciones a las teorías, y deja abierta la posibilidad de que podamos hacer inferencias *falsadoras:* una inferencia de la verdad del enunciado de una observación («esto es un cisne negro») a la falsedad de una teoría («todos los cisnes son blancos») puede ser perfectamente válida deductivamente.

No veo como puede cuestionarse esta solución a la formulación desafiadora de Russell (excepto confundiendo la primera fase del problema con la segunda o la tercera fase, que falta por exponer). Parece que Russell, al sugerir que la ciencia sería imposible si Hume estuviera en lo cierto, simplemente pasó por alto el importantísimo hecho de que el argumento de Hume no muestra la invalidez de las inferencias falsadoras, de la observación a la teoría. La sugerencia de Russell encaja en la concepción de Kant de la ciencia como conocimiento bien establecido *(scientia, episteme),* pero no encaja en la concepción del propio Russell de las teorías científicas como hipotéticas o conjeturales. De hecho, en muchos pasajes, Russell describe el método de la ciencia de un modo que hace innecesario hablar

[33] Bertrand Russell, *History of Western Philosophy,* 1945. Véase el penúltimo párrafo del capítulo XVII (p. 673 de la primera edición inglesa). El pasaje original no está subrayado.

de la inducción; escribe por ejemplo: «La lógica, en lugar de ser, como antes, la barrera de las posibilidades, ha venido a ser la gran liberadora de la imaginación, a presentar innumerables alternativas que le están cerradas al sentido común irreflexivo, *y a dejar a la experiencia la tarea de decidir, cuando una decisión es posible, entre los muchos mundos que la lógica ofrece a nuestra elección*»[34]. Las palabras clave en este bello pasaje son quizá, «cuando una decisión es posible». No me cabe duda de que al escribir este pasaje, Russell vio que, cuando una decisión es posible, sólo puede serlo como rechazo de algunas de las «innumerables alternativas», más que como una decisión positiva en favor de una de ellas. Pero aparentemente no vio que éste era el caso típico o incluso el único; o que este hecho hace posible la resolución del problema lógico de la inducción, de Hume.

Sólo hay que añadir un punto importante a esta solución del problema: la observación sola no siempre es capaz de decidir cuál de las dos teorías en competencia es mejor, aunque *puede* hacerlo, especialmente si las teorías permiten un experemiento crucial. En general, se necesita más que observación: también se necesita una discusión crítica de los méritos de las dos teorías. Tal discusión debe tener en cuenta si resuelven el problema que se supone que tienen que resolver; si explican lo que se supone que tiene que explicar; si no se reducen a cambiar el problema, por medio de un supuesto *ad hoc* incontrastable, por ejemplo; si son contrastables y qué grado de bondad pueden alcanzar sus contrastaciones.

Estas cuestiones (estrechamente relacionadas con lo que llamo «el problema de demarcación» que se discutirá en el capítulo 2, secciones 17 a 26) son muy importantes. Porque puede ocurrir que nos enfrentemos a dos teorías de un valor muy diferente y que nuestras observaciones sean igualmente compatibles con las dos, aunque por razones muy diferentes: con una teoría, a pesar de que la someten a una contrastación rigurosa, y con la otra, simplemente, porque no es contrastable, porque todas las observaciones son compatibles con ella. (La primera teoría puede ser la de Newton, o digamos, la teoría de Kepler de que todos los planetas se mueven en elipses, combinada quizá con su teoría de la Armonía Universal; y la segunda puede ser la teoría de Platón de que todos los planetas tienen alma y son dioses.)

Sólo una vez que se han resuelto las cuestiones respecto al valor explicativo y la contrastabilidad de las dos teorías, pode-

[34] Bertrand Russel, *The Problems of Philosophy*, 1912, p. 148. (Sin subrayar en el original.)

mos decir de ellas si realmente compiten entre sí y si pueden ser sometidas a contrastaciones cruciales basadas en observaciones, que puedan decidir en contra de una de ellas y mostrar así que la otra es «mejor». De este modo podemos llegar a decir, al final, después de muchas pruebas y muchos errores, que tenemos una teoría que, según el estado actual de nuestra discusión crítica, incluidas las contrastaciones en forma de observaciones, parece acercarse más a la verdad que todas las demás que están bajo consideración.

II

Así, hemos hecho frente al desafío de Russell y hemos tratado de la primera fase del problema de la inducción.

El problema de «la creencia racional», como puede llamarse a la segunda fase de nuestro problema, es, en mi opinión, menos fundamental e interesante que la primera. Surge como sigue:

Incluso aunque admitamos que no hay dificultad lógica en mostrar que las observaciones pueden ayudarnos a veces a distinguir entre teorías «buenas» y «malas» y cómo pueden hacerlo, debemos insistir en que no se ha dado explicación alguna sobre la fiabilidad de la ciencia, o sobre el hecho de que es *razonable creer* en sus resultados, en teorías bien contrastadas por medio de observaciones. Una buena teoría no lo es sólo porque hasta el momento haya escapado a la falsación: incluso aunque admitamos que somos siempre falibles y muy propensos a cometer errores, y que todas las teorías científicas son conjeturales, no es razonable negar que hay una cantidad enorme de conocimiento positivo en la ciencia. Pero ¿cómo podemos admitir lo razonable de esta posición —de este creer, bastante restringido, en la ciencia— y admitir al mismo tiempo que Hume tiene razón?

Este es el nuevo desafío. Mi concepción de que es menos fundamental e interesante que el primero se debe en parte a mi limitado interés por la filosofía de la creencia. Pero también se debe al hecho de que no se necesitan ideas realmente nuevas para enfrentarse a él. No obstante, este nuevo desafío, o la segunda etapa del problema, puede ayudar a clarificar la situación.

Supondré que se ha llegado a un acuerdo sobre el carácter conjetural de las teorías científicas: que nuestras teorías científicas siguien siendo inciertas por mucho «éxito» que tengan y muy bien «fundamentadas» que estén por la evidencia y el resultado de la discusión y por muy incapaces que seamos de

prever el tipo de cambio que será necesario. (¡Recuérdese la mecánica de Newton!) En consecuencia, supondré que si hablamos aquí de «creencia racional» en la ciencia y en la teoría científica *no* queremos decir que es racional creer en la verdad de ninguna teoría en particular. Este punto es de la mayor importancia.

¿Cuál es, entonces, el objeto de nuestra «creencia racional»? Es, propongo, no la verdad, sino lo que podemos llamar *el parecido con la verdad* (o la «*similitud con la verdad*») de las teorías de la ciencia, en la medida en que han resistido una crítica rigurosa, incluidas contrastaciones. Lo que creemos (con razón o sin ella) no es que la teoría de Newton o la teoría de Einstein sean verdaderas, sino que son *buenas aproximaciones* a la verdad, aunque susceptibles de ser sustituidas por otras mejores.

Pero esta creencia, afirmo yo, *es* racional. Es racional incluso si suponemos que mañana vamos a descubrir que las leyes de la mecánica (o las que nosotros considerábamos leyes de la mecánica) han cambiado repentinamente (una posibilidad que discutiremos más extensamente en las subsecciones III y IV, dedicadas al «problema del mañana»). Porque en este caso nos enfrentaríamos al problema de explicar, no sólo las nuevas regularidades observadas, sino también las antiguas. Nuestro problema sería: *a)* construir una teoría de la cual pudiera obtenerse la antigua, bajo ciertas condiciones, como una *buena aproximación*, y *b)* mostrar qué circunstancias (condiciones iniciales) produjeron el cambio[35]. Este enfoque, que asegura la supervivencia, como aproximación de la teoría sustituida, lo exigen el realismo y el método de la ciencia. Someterse simplemente al hecho de que ha ocurrido el cambio y reducirse a describirlo equivaldría a la aceptación de *milagros*, al abandono de la búsqueda de explicación racional y, por tanto, al abandono de la tarea de la ciencia: la búsqueda de racionalidad.

Estas consideraciones muestran que la creencia en el parecido con la verdad de los resultados bien corroborados de la ciencia (tales como las leyes de la mecánica) es ciertamente racional y sigue siéndolo aun después de que esos resultados hayan sido sustituidos. Además es una creencia que admite *grados*.

Tenemos que distinguir entre dos dimensiones diferentes o dos escalas de grados: el grado de parecido con la verdad de una teoría y el grado de racionalidad de nuestra creencia de que una cierta teoría ha logrado (un cierto grado de) parecido con la verdad.

---

[35] Véase *L. I. C.*, sección 79.

He llamado al primero de estos dos grados «grado de similitud con la verdad»[36] y al segundo, «grado de corroboración»[37]. Son «comparativos» en el sentido de que dos teorías pueden ser comparadas respecto a la similitud con la verdad o a la corroboración sin llevar, en general (esto es, con la posible excepción de teorías probabilísticas), a evaluaciones numéricas.

Si dos teorías rivales han sido criticadas y contrastadas tan completamente como podamos hacerlo, con el resultado de que el grado de corroboración de una de ellas es mayor que el de la otra, en general *tendremos razón* para creer que la primera es una mejor aproximación a la verdad que la segunda. (También es posible decir que una teoría que aún no ha sido corroborada, que es *potencialmente* mejor que otra; es decir, que sería razonable aceptarla como una mejor aproximación a la verdad, siempre que pasase ciertas contrastaciones)[38].

según esta concepción, la racionalidad de la ciencia y de sus resultados —y, por tanto, de la «creencia» en ellos— está esencialmente ligada a su progreso, con la siempre renovada discusión sobre los méritos relativos de las teorías nuevas; está ligada al *derrocamiento progresivo* de las teorías más que a su supuesta *consolidación progresiva* (o creciente probabilidad), resultante de la acumulación de observaciones en su apoyo, como creen los inductivistas.

Un ejemplo, entre cientos, puede ilustrar lo que digo.

La mayoría de nosotros creemos firmemente, hoy en día —y tenemos razón para creer—, en el modelo copernicano del sistema solar (tal como lo revisaron Kepler y Newton): en una cierta disposición de los planetas, que se mueven en elipses vecinas, aocmpañados por sus lunas. Pero ¿cuáles son nuestras razones para creer en el parecido de esta teoría con la verdad? (Desde luego, no creemos en su verdad completa, puesto que no es más que un modelo y, por tanto, es inevitablemente una ultra-simplificación y una aproximación totalmente independiente del hecho de que pueda estar necesitado de una corrección einsteiniana y quizá de mayor elaboración revolucionaria para dar razón de la validez aproximada de la ley de Bode.)

Contar *la historia de la evidencia observacional* que se ha acumulado durante muchos siglos, empezando con los egipcios

---

[36] Véase *El desarrollo del conocimiento científico (C. & R.)*, cap. 10 y Addenda. [Véase también *Conocimiento objetivo (Obj. Kn.)*, capítulos 2, 3 y 9; y «A Note on Verisimilitude», en *British Journal for the Philosophy of Science* 27, 1976, pp. 147-64. *Ed.*]

[37] Véase *L. I. C.*, Capítulo X y Apéndice *IX.

[38] *Cf.* con esto el capítulo 10 de *El desarrollo del conocimiento científico (C. & R.)*.

y los babilonios, no haría justicia a las poderosas razones que tenemos para creer en el parecido con la verdad, del modelo. Independientemente del hecho de que tendríamos que mencionar tradiciones (como el relato de Homero sobre la vacilación del sol en su camino o el de Josué sobre su detención) que tendrían que explicarse (por medio de la discusión crítica) como mitos, los informes sobre observaciones son sumamente selectivos y esta selección está influenciada por ideas preconcebidas.

Nuestras razones para ceer en el modelo copernicano son mucho más fuertes. Consisten en *la historia de la discusión crítica,* incluyendo la evaluación crítica de las observaciones, de todas las teorías del sistema solar desde Anaximandro, sin descuidar la hipótesis de Heráclito de que cada día nacía un nuevo sol, o las cosmologías de Demócrito, Platón, Aristóteles, Aristarco y Ptolomeo. No fue tanto la acumulación de observaciones por parte de Tycho como el rechazo crítico de muchas conjeturas por parte de Kepler, Descartes y otros, que culminó en la mecánica de Newton, y su posterior examen crítico, lo que persuadió finalmente a todo el mundo de que se había dado un gran paso hacia la verdad.

Esta convicción, esta creencia, esta preferencia, es razonable, porque está basada en el resultado del estado actual de la discusión crítica; y una preferencia por una teoría puede llamarse «razonable» si es argüible y si resiste un *argumento crítico penetrante:* si resiste a los ingeniosos intentos de mostrar que no es verdad, o que no está más cerca de la verdad que sus competidoras. En realidad, *éste es el mejor significado de «razonable» que conozco.*

Lo razonable de una creencia, en el sentido que se ha descrito aquí, cambia con el tiempo y con las tradiciones culturales y, en un grado limitado, incluso con el grupo de personas que participan en la discusión; porque nuevos argumentos, nuevas ideas críticas, pueden alterar lo razonable de una creencia. No necesito decir que nuevos experimentos pueden hacer lo mismo.

Sin embargo, *el estado actual de la discusión crítica puede ser muy definido respecto a la superioridad de una teoría sobre otra. Puede ser también muy definido respecto a la falsedad de una teoría, pero no respecto a su verdad.* Cuando esas valoraciones tan definidas han sido el resultado de una discusión crítica prolongada y completa en el pasado, entonces, habitualmente han sido confirmadas por discusiones posteriores; han sido notoriamente infrecuentes las inversiones de las evaluaciones relativas de dos teorías y también los renacimientos de teorías rechazadas definitivamente. (La teoría corpuscular de la luz de Newton no fue resucitada en modo alguno por la teoría del fotón de Einstein, como se afirma con frecuencia: era una teo-

ría de la transmisión o propagación; la teoría del fotón, concebida para hacer frente a los problemas de la emisión y la absorción, sucumbe ante la teoría ondulatoria en los problemas de propagación.) Contrariamente a esto, las inversiones de afirmaciones basadas en «evidencia inductiva» han sido sorprendentemente frecuentes, de modo que hay algo parecido a evidencia inductiva contra la inducción. (No sólo el falso enunciado de que todos los cisnes son blancos, sino también la mayoría de las supersticiones, incluidas las supersticiones médicas, se apoyan en enormes cantidades de evidencia inductiva.) El hecho de que las inversiones de evaluaciones críticas sean escasas puede explicarse fácilmente: los buenos argumentos críticos retienen su poder, a menos, naturalmente, que operen con prejuicios no detectados o con evidencia basada en observaciones ficticias o mal interpretadas.

Decir que han sobrevivido muchas de nuestras evaluaciones críticas y que, incluso cuando nuestras mejores teorías han sido reemplazadas, rara vez hemos invertido el juicio de que eran las mejores de que se disponía en el momento en que se las juzgó así, es decir que nuestro método crítico ha sido sorprendentemente acertado en el pasado. *Pero no debemos concluir que seguirá siéndolo en el futuro.* Nuestros problemas pueden llegar a ser demasiado difíciles para nosotros, o nuestro intelecto puede entrar en decadencia. Después de todo, sólo unos cuantos entre los miles de científicos bien formados consiguen hacer contribuciones a los problemas más difíciles y más fundamentales de la ciencia; y si esos pocos dejaran de estar disponibles, la ciencia podría estancarse. O algunos prejuicios pueden convertirse en nuestra ruina: el culto a los tecnicismos impresionantes o el culto a la precisión pueden vencernos o interferir con nuestra búsqueda de claridad, simplicidad y verdad. No hay un «camino real» hacia la ciencia; no hay método que garantice el éxito; y cualquier teoría del conocimiento que, al explicar por qué tenemos éxito, nos prediga que continuaremos teniéndolo, explica y predice demasiado.

Esta selección de la mejor teoría es como la del mejor testigo. Cuando nos encontramos con testigos que se contradicen mutuamente tratamos de repreguntarles, de analizar críticamente lo que dicen, de comprobar por segunda vez los detalles pertinentes. Y podemos decidir —decidir racionalmente— preferir a uno de ellos: aunque supongamos que todos los testigos, incluidos los mejores, son algo parciales, puesto que todo testimonio, incluso el que se limita a la observación, es selectivo (como todo pensamiento), de modo que el ideal de «toda la verdad y nada más que la verdad» es, estrictamente hablando, inalcazable; aunque, sin duda, estamos dispuestos a reconside-

rar nuestras preferencias a la luz de nuevos argumentos críticos o de nueva evidencia.

Esta es, en esquema, la contribución «positiva» que hace mi teoría a la segunda fase del problema de la inducción. Es necesario, sin embargo, hacer hincapié también en su aspecto negativo.

Aunque podemos creer razonablemente que el modelo copernicano, tal como lo revisó Newton, está más cerca de la verdad que el de Ptolomeo, no hay forma de decir *cómo* está de cerca: incluso aunque pudiéramos *definir* un metro para la similitud con la verdad (lo que sólo podemos hacer en casos que parecen tener escaso interés) seríamos incapaces de aplicarlo a menos que conociéramos la verdad, lo que no es el caso. Podemos pensar que nuestras ideas actuales sobre el sistema solar están cerca de la verdad y puede que así sea; pero no podemos saberlo. Tampoco debemos pensar que nuestra discusión nos ha proporcionado razones para creerlo: todo lo que ha hecho ha sido proporcionarnos buenas razones para creer que hemos progresado hacia la verdad: es decir, que tenemos buenas razones para creer que algunas de nuestras ideas son *más parecidas a la verdad* que algunas de sus alternativas. Y aunque el modelo copernicano fue *el* gran punto de ruptura, ya no pensamos que nuestro sol es el centro del universo, ni siquiera de nuestra galaxia.

No podemos justificar nuestras teorías, ni la creencia de que son verdaderas; tampoco podemos justificar la creencia de que están cerca de la verdad. Podemos, no obstante, defender racionalmente una preferencia —a veces una preferencia muy fuerte— por una cierta teoría, a la luz de los resultados actuales de nuestra discusión.

El método de la ciencia es racional[39]: es el mejor que tenemos. Es, por tanto, racional el aceptar sus resultados; pero no en el sentido de hacer depender nuestra fe de ellos; nunca sabemos de antemano dónde pueden fallarnos.

Y, sin embargo, *es* razonable, o racional, fiarse de los resultados de la ciencia *para todos los fines prácticos*. Porque la práctica siempre implica una elección: podemos actuar de esta forma o de esta otra. (La inacción no es, naturalmente, más que una posible forma de acción.) Y en la medida en que aceptamos o rechazamos una teoría científica como fundamento de

[39] R. A. Wollheim, en una reseña de *L. I. C., (The Observer,* 15 febrero 1959), expresa este punto con mucha claridad: «...mientras pensemos que los científicos están siempre tratando de establecer leyes, su actividad nos parecerá irracional; pero una vez que nos los imaginamos tratando de falsar hipótesis continuamente, entonces, la tarea que les ocupa parece consecuente y comprensible».

la acción práctica, en lugar de otra, esto significa escoger una teoría en vez de otra. Cuando estemos en la situación de poder elegir, será racional elegir, entre dos teorías en competencia, la que haya superado una prolongada discusión crítica, contrastaciones incluidas.

Un último punto debe ser expuesto aquí. La creencia parece ser algo muy necesario en las acciones prácticas: el hombre es un animal que cree porque es un animal que actúa. El teórico, en cuanto teórico, puede pasarse sin creencia. Para él, la teoría que parece tener la mayor similitud con la verdad no es una en la que creer, sino una que es importante para continuar progresando. Es también la que él designará como digna de seguir siendo criticada.

Naturalmente, incluso el teórico, en cuanto teórico, tiene que actuar; por ejemplo, tiene que escoger sus problemas. Y, en la medida en que lo hace, también él puede ir guiado por creencias y por dudas.

III

La tercera fase del problema de la inducción es, en mi opinión, menos importante aun que la segunda. La primera fase era un problema práctico urgente sobre el método: cómo distinguir entre teorías buenas y malas. La segunda era menos importante, pero seguía siendo de cierta urgencia. Porque creemos, en efecto, en los resultados de la ciencia, en un sentido corriente de «creer» y esta creencia es razonable en un sentido corriente de «razonable». Al no ser inmediatamente obvio cómo podría darse razón de estos hechos dentro del marco lógico de nuestra solución, teníamos aquí un problema de cierta significación.

Pero la tercera fase del problema, el «problema del mañana», como lo llamaré, me parece a mí, una vez que hemos alcanzado la claridad de las dos primeras fases, que no es más que el típico embrollo filosófico (a menos, claro, que se confunda con la cuarta fase, metafísica). Admitamos que un inductivista como Hume o como Russell puede pensar que es indistinguible de lo que he llamado el problema de la inducción y puede incluso pensar que supone una formulación superior del mismo problema. Pero una parte esencial de mi solución no-inductivista consiste precisamente en reconocer el carácter fundamental de la primera fase del problema y el carácter inferior de su tercera fase. Podemos formular la tercera fase del problema diciendo: «¿Cómo sabe usted que el futuro será igual que

el pasado?» o quizá «¿cómo sabe usted que las leyes de la naturaleza seguirán siendo válidas mañana?».

La respuesta más simple y más directa a la primera de esas dos preguntas es: «Yo *no* sé que el futuro vaya a ser igual que el pasado; por el contrario, tengo buenas razones para esperar que será diferente en muchos aspectos, de hecho en todos aquellos aspectos que los inductivistas mencionan como ejemplos de 'la uniformidad de la naturaleza'. Así, nuestro acostumbrado pan de cada día puede convertirse en veneno (recuérdese el caso francés de envenenamiento en masa con el cornezuelo del centeno); el aire puede afixiar a quienes lo respiran (recuérdense los envenenamientos del aire en Hamburgo); y nuestros mejores amigos, en los que más confianza tenemos, pueden convertirse en nuestros enemigos mortales (recuérdense las sociedades totalitarias)»[40].

Si se nos dice que esta respuesta a la primera pregunta está fuera de lugar porque lo que significaba está más claramente enunciado en la segunda pregunta («¿Cómo sabe usted que las leyes de la naturaleza seguirán siendo válidas mañana?»), entonces mi respuesta vuelve a ser enfática: «Yo *no* sé que lo que hoy consideramos una ley de la naturaleza no vaya a considerarse mañana una conjetura refutada. De hecho, esto parece ocurrir con más frecuencia que los envenenamientos con el cornezuelo del centeno».

Pero pueden decirme que, de nuevo, esta respuesta está fuera de lugar, que la pregunta no era sobre hipótesis que pueden refutarse, sino sobre las auténticas y verdaderas leyes de la naturaleza, auténticas regularidades naturales posibilidad de que éstas cambien. Mi respuesta a esta pregunta es trivial. Hay todo tipo de cambios en la naturaleza; pero lo que llamamos una ley de la naturaleza es un enunciado de algo que permanece *invariable* durante los cambios; y si descubrimos que lo que habíamos creído que era invariable cambia, entonces habíamos hecho una conjetura equivocada: simplemente, no era una ley de la naturaleza.

He descubierto, sin embargo, que los inductivistas no quedan satisfechos con estas respuestas y que no piensan que la tercera fase del problema, el problema del mañana, esté resuelta.

Sus recelos están ligados a algunas de mis concepciones referentes al «grado de corroboración» (véase el capítulo IV). A menudo lo he descrito como *nada más* que un informe resu-

[40] Mis ejemplos aluden a los que Russel dio con la intención contraria —la de defender la inducción— en su obra *Problems of Philosophy*, p. 69. [Vease *Conocimiento objetivo* (*Obj. Kn.*), cap. I. *Ed.*]

mido, o una valoración del modo en que la teoría ha resitido, hasta el momento, las críticas y las contrastaciones, así como de la minuciosidad de la discusión crítica y la severidad de las contrastaciones a las que se la ha sometido.

Algunas veces lo he descrito también como el grado en el que la teoría en cuestión «ha sido capaz de mostrar su aptitud para sobrevivir mediante la superación de las contrastaciones»[41]. Pero como muestra el contexto de esos pasajes, no quería decir con esto más que un informe sobre la pasada aptitud de la teoría para sobrevivir las contrastaciones severas: como Darwin, no supuse que algo (fuese animal o teoría) que hubiera demostrado aptitud para sobrevivir a las contrastaciones sobreviviéndolas, hubiese mostrado su aptitud para sobrevivir todas o la mayoría de las contrastaciones futuras[42]. De hecho, creo que una teoría, por muy bien contrastada que esté, puede ser mañana, sobre todo si alguien pone todo su empeño en refutarla y sobre todo si ese alguien tiene una nueva idea sobre cómo contrastarla.

Pero si el grado de corroboración no es nada más que una evaluación de la actuación de la teoría en el pasado ¿no surge de nuevo el problema de la inducción, bajo la forma del problema de mañana? Porque el grado de corroboración de una teoría, es decir, su actuación pasada, ¿no determina nuestras expectativas respecto a sus actuaciones futuras? ¿No estoy yo mismo, a pesar de mi negativa, atribuyendo erróneamente a una teoría la disposición de sobrevivir contrastaciones futuras basándome en su actuación pasada?

Estoy de acuerdo en que tal atribución por mi parte equivaldría al desbaratamiento de mi teoría: sería una inferencia inductiva. Pero no necesito ir más allá de los argumentos ya avanzados[43] para aclarar la cuestión.

La cuestión es ésta: yo no creo que una teoría corroborada en alto grado tenga mayores posibilidades, o probabilidades o lo que sea, de sobrevivir a futuras contrastaciones; ni que sea más probable que lo haga ella que una teoría corroborada en un grado menos alto. Por el contrario, la probabilidad de que sobreviva una teoría dependerá en gran medida, entre otros factores, del ritmo del progreso en esa rama de la ciencia o,

---

[41] *L. I. C.*, cap. X, antes de la sección 79 (p. 234).

[42] El argumento sobre «la aptitud para sobrevivir» lo entendió mal mi reseñista de *Mind* en el sentido que se indica aquí; *cf.* nota 18 a la sección 2 más arriba.

[43] Esta discusión desarrolla varios pasajes de *L. I. C.*, sobre todo la sección 79 y el Apéndice *IX (subsección *14), y de «Philosophy of Science: A Personal Report» (1957; ahora capítulo 1 de *El desarrollo del conocimiento científico* [*C. & R.*]).

dicho de otro modo, del interés que se tomen los investigadores en ese terreno determinado y de sus esfuerzos por concebir nuevas contrastaciones. Pero el ritmo del progreso puede ser muy alto justo en un terreno en el que los criterios para la crítica y para la contrastación sean muy rigurosos, es decir, en terrenos en los que tenemos teorías con un alto grado de corroboración.

En consecuencia, yo *no* debería esperar que una teoría muy corroborada fuera a sobrevivir, en general, a una teoría menos bien corroborada. La esperanza de vida de una teoría no aumenta, creo yo, con su grado de corroboración ni con su pasada capacidad para sobrevivir las contrastaciones.

Pero (se me preguntará) ¿no espero yo que el sol salga mañana, o no baso mis predicciones en las leyes del movimiento? Naturalmente, lo hago, porque son las mejores leyes de que disponemos, como he expuesto largamente más arriba. Incluso en donde tengo dudas teóricas basaré mis acciones (si tengo que actuar, es decir, escoger) en la elección de la mejor teoría disponible. Así, estaría dispuesto a apostar a que el sol saldrá mañana (apostar es una acción práctica), pero no estoy dispuesto a apostar a que las leyes de la mecánica de Newton (o de Einstein) van a sobrevivir a futuras críticas o a sobrevivirlas durante más tiempo que, digamos, la mejor teoría disponible sobre la transmisión sináptica, a pesar de que la última tiene (o eso parece) un menor grado de corroboración. En cuanto a acciones prácticas (tales como apuestas sobre predicciones hechas por esas teorías), estoy dispuesto a basarlas, en los dos casos, en la mejor teoría en su terreno, siempre que haya sido bien contrastada.

Esto puede expresarse también así. La cuestión de la superviviencia de una teoría es una cuestión perteneciente a su destino histórico y, por tanto, a la historia de la ciencia. Por otro lado, su uso para la predicción es una cuestión relacionada con su aplicación. Esas dos cuestiones están relacionadas, pero no íntimamente. Porque con frecuencia aplicamos las teorías sin ninguna vaciliación, incluso aunque estén muertas —es decir, falsadas—, siempre que sean aproximaciones suficientemente buenas para nuestro propósito. Así pues, no hay nada de paradójico en que yo esté dispuesto a apostar por las aplicaciones de una teoría y al mismo tiempo me niege a apostar por la superviviencia de la misma teoría.

Mi negativa a apostar por la superviviencia de una teoría bien corroborada muestra que no extraigo ninguna conclusión inductiva de la supervivencia pasada a la supervivencia futura.

IV

Pero ¿es cierto que no saco conclusiones inductivas de la actuación pasada a la actuación futura? ¿No tiene razón Russell en resaltar que «la *única* razón para creer que las leyes del movimiento seguirán operando es que han operado hasta ahora, en la medida en que nuestro conocimiento del pasado nos permite juzgar»?[44]. Esta concepción inductivista, a pesar de su capacidad de persuasión, es equivocada.

Primero debemos recordar que nadie ha observado jamás las leyes del movimiento en operación: las leyes del movimiento son un invento nuestro y se han inventado para resolver ciertos problemas, para explicar ciertos sucesos. Si han tenido éxito en ello, entonces confiaremos en ellas en vez de en sus rivales de menor éxito. Pero mañana alguien puede presentarnos una nueva teoría, un nuevo conjunto de leyes del movimiento que no sólo resuelva todos los problemas y pase todas las contrastaciones que pasó y resolvió el antiguo conjunto de leyes, sino que sugiera además nuevas contrastaciones cruciales. Las nuevas leyes pueden sugerirnos desviaciones de las leyes antiguas que nos habían pasado desapercibidas hasta ahora. (Todo esto ha ocurrido más de una vez desde la publicación del libro de Russell, de hecho, una vez antes de que se cumpliera el año de su aparición. Estoy suponiendo aquí que Russell se refería con «las leyes del movimiento» a un conjunto de leyes invariables de Lorenz, puesto que, de otro modo, habría ocurrido incluso antes de que él lo escribiera.) Así pues, estas leyes del movimiento que «han operado hasta ahora en la medida en que *nuestro conocimiento* del pasado nos permite juzgar» *no* «habían estado operando» y *no* había razón para esperar que lo hicieran, porque, en la medida en que nuestro conocimimiento actual nos permite juzgar, nunca estuvieron «en operación».

No obstante, si en lo que estaba pensando Russell era en aquellas otras leyes del movimiento, las leyes del movimiento *reales* que, podemos suponer, operaron en el pasado, entonces tenemos que resaltar que nos eran (y nos siguen siendo) *desconocidas* y que ninguna inferencia inductiva —ninguna inferencia de su operación pasada a su operación futura— puede basarse en esas leyes *desconocidas*.

Naturalmente, si unas leyes *reales* han operado en el pasado (como yo creo), continuarán operando en el futuro. Pero *esta* afirmación no está basada en la inducción: está basada en el

[44] Bertrand Russell, *The Problems of Philosophy*, p. 61. El subrayado es de Russell.

hecho de que explicamos los cambios valiéndonos de leyes invariables y que debemos negarnos a llamar ley verdadera (o real) a nada que no «opere» —es decir, que no sea válido—en todo momento y lugar.

Podría sugerirse que interpretásemos las palabras de Russell de una forma más libre, con el sentido de «la *única* razón para creer que se observará al sol salir mañana es que se le ha observado salir en el pasado; y la única razón para creer que la tierra continuará rotando es, si nuestra interpretación de nuestra observación es correcta, que lo ha hecho en el pasado. (Y comentarios análogos serán válidos para las leyes del movimiento en general.)» Pero está claro que esto tampoco funciona (con total independencia del hecho de que las observaciones —que, como hemos dicho, son selectivas— no sean la *única* razón de cualquier creencia razonable o racional). Porque decir que nuestra interpretación de movimientos pasados es *correcta* es decir que ciertas teorías son verdaderas.

No es la formulación de Russell la que es equivocada, sino los sentimientos que expresa (incluidos sus diversas referencias a la probabilidad). Simplemente *no* hay razón para creer en la *verdad* (o la probabilidad) de ningún conjunto determinado de conjeturas al que llamamos teoría física; aunque puede haber razones para preferir una teoría a otras por ser una mejor *aproximación a la verdad* (que no es una probabilidad)[45]. Esto es totalmente distinto.

V

Los inductivistas aplican un método justo y razonable de argumentación si tratan de mostrar que esas dificultades, de sobra conocidas, con las que han luchado durante tanto tiempo y con tan poco éxito, deben surgir, quizá con ciertos cambios apropiados, también dentro de mi teoría.

Una de las dificultades que tienen los inductivistas, me parece a mí, consiste en decir por qué, si sólo se han observado cuervos negros hasta digamos 1950, debe uno preferir la ley «todos los cuervos son negros» a la ley «todos los cuervos antes de 1970 son negros y después de 1970 serán blancos». Porque esas dos leyes parecen explicar igualmente bien la evidencia proporcionada por las observaciones. No obstante, se dice, pre-

[45] Es decir, la similitud con la verdad no cumple las reglas (por ejemplo las reglas de Keynes) del cálculo de probabilidades. Véase *El desarrollo del conocimiento científico (C. & R.)*, capítulo 10 y *Addenda*.

ferimos «obviamente» la primera a la segunda. El problema consiste en explicar el porqué de esta preferencia[46].

Ahora bien, puede pensarse que el mismo problema debe surgir dentro de mi teoría sobre la resolución de problemas por medio de conjeturas, como sigue. Alguien puede sorprenderse por el hecho de que los dos o tres cuervos que ha visto sean negros; quiere una explicación y da en la hipótesis de que todos los cuervos son negros. Pero otra presona sugiere la hipótesis de que todos los cuervos son negos hasta 1970; o, para evitar la objeción de que nuestras leyes deben contener sólo universales (y «1970» es un nombre individual[47]), sugiere una tercera hipótesis: que todos los cuervos son negros excepto los fotografiados durante un eclipse total de sol (que son blancos).

¿Por qué, se me pregunta, preferimos la primera hipótesis a la segunda y a la tercera, a pesar que las tres parecen concordar igualmente con la evidencia?

La respuesta es que la segunda y la tercera teorías afirman una conexión, una dependencia, que *tendría que ser explicada:* porque aceptar un cambio inexplicado e inexplicable en una determinada fecha equivaldría a aceptar un milagro[48]. (Véase la subsección II más arriba.) Esto, por sí solo, las hace inferiores a la primera. Además, nuestras teorías actuales respecto al color de los pájaros son de un carácter que nos hace sospechar que cualquier explicación del cambio en cuestión que no fuese *ad hoc* entraría en conflicto con algunas teorías bien corroboradas (de la genética). Por esas solas razones, no tengo la menor dificultad para explicar mi preferencia en un caso como éste.

---

[46] No estoy seguro de cuándo se escribió este pasaje, pero es casi seguro que ocurrió antes de que el Prof. Nelson Goodman me llamase la atención sobre su libro *Fact, Fiction and Forecast,* 1955; véase la nota 72 al cap. 11 de *El desarrollo del conocimiento científico (C. & R.).* Sea como fuere, mi discusión muestra que el problema del apoyo inductivo planteado por Goodman con ayuda de sus nuevos y peculiares predicados («horripil» [grue], etc.), puede enunciarse fácilmente sin tales predicados. Así, el problema de esos nuevos predicados y el problema de su exclusión no me parecen un nuevo enigma que haya que resolver por medio de una teoría de atrincheramiento, sino más bien una pista falsa (o quizá «equivocada») [a red (or perhaps a reen) herring]. [Véanse mis artículos, «Goodman's Paradox: A Simple-Minded Solution», en *Philosophical Studies,* diciembre 1968, pp. 85-88; «Theories of Demarcation between Science and Metaphysics», en I. Lakatos y A. Musgrave, eds., *Problems in the Philosophy of Science* (Amsterdam: North-Holland Publishing Company, 1968), pp. 54-57; «Eine Lösung des Goodman-Paradoxons», en Gerard Radnitzky y Gunnar Andersson, eds., *Voraussetzungen und Grenzen der Wissenschaft* (Tübingen: J. C. B. Mohr (Paul Siebeck) Verlag, 1981); y «The Philosophy of Karl Popper: Part III, Rationality, Criticism and Logic» *Philosophia,* 1982. *Ed.*]

[47] *Cf. L. I. C.,* secciones 13 a 15 y Apéndice *X.

[48] [*Cf. L. I. C.,* sección 79. *Ed.*]

Sin embargo, mi principal objeción al argumento es bastante distinta.

Todas nuestras hipótesis son conjeturas y cualquiera es libre de ofrecer conjeturas, incluso conjeturas que nos pueden parecer bastante tontas a la mayoría de nosotros. Sólo así podemos abrir camino a nuevas ideas, atrevidas y no convencionales. Tenemos que pagar el precio de esta libertad encontrándonos con frecuencia con ideas que parecen tontas. Pocas de éstas se tomarán en serio; pero puede que algunas se tomen en serio; y algunas pueden, en ocasiones, contradiciendo las primeras impresiones, resultar movimientos en la buena dirección. Por tanto, aunque todo científico que desecha *a priori* una teoría por tonta corre un riesgo, no hay forma de evitar esos riesgos; no sólo es arriesgada toda propuesta de una nueva conjetura, sino que también lo es la decisión de tomarla en serio o de descartarla por absurda. Al contrario que los inductivistas, yo no afirmo que hay *una* teoría (alcanzada inductivamente) que dé razón y explique mejor que ninguna cualquier evidencia dada. Por el contrario, la idea de *una pluralidad de conjeturas en competencia* —que reconozco que tratamos de reducir por medio de la crítica— es esencial para mi metodología.

Mientras tal pluralidad puede desconcertar al inductivista, a mí no me crea ningún problema.

En la práctica, muchas de las conjeturas más obviamente «tontas» pueden eliminarse por medio de la crítica; por ser incontrastables o menos contrastables; por ser arbitrarias o *ad hoc*; por crear, sin excusa, nuevos problemas innecesarios y por chocar con nuestras ideas más generales de lo que debe ser una explicación satisfactoria: ideas que son algo vagas, pero que, como las teorías científicas, se desarrollan por el método de prueba y error; ideas del «estilo» (mecánico, estadístico, etc.) de una buena explicación. Naturalmente, tales consideraciones pueden llevarnos a rechazar una *buena* teoría. Este también es uno de los riesgos que corremos; forma parte del carácter conjetural de la ciencia.

Así, Galileo rechazó equivocadamente la teoría lunar sobre las mareas porque (como he sugerido en otro lugar)[49] formaba parte de la teoría astrológica de las «influencias» de los astros, que él consideraba, con razón, de mal gusto, aunque nadie parece haberse tomado nunca la molestia de refutar la astrología utilizando la evidencia proporcionada por la observación, ni de examinar críticamente la enorme cantidad de evidencia inductiva en la que se «basa». Otros ejemplos son el rechazo de

[49] Véase mi *El desarrollo del conocimiento científico (C. & R.)*, p. 188, nota 4 (y la referencia que allí se hace a la p. 38, nota 4).

la teoría de la gravitación de Newton, al principio, en el continente y, en Inglaterra, bajo la influencia de la autoridad de Newton, el rechazo, al principio, de la teoría ondulatoria de la luz de Huygens, hasta que Young refutó la teoría corpuscular.

Por tanto, es ingenuo creer que es deber del metodólogo dar cuenta de la tontería de las teorías tontas que encajan con los hechos y dar razones para su exclusión *a priori*: debemos dejar a los científicos la tarea de luchar por el reconocimiento y la supervivencia de sus teorías (y, por los suyos propios). Además, es una creencia inductivista. El inductivista que piensa que la única razón suficiente para aceptar una teoría es el apoyo de observaciones pasadas, queda desconcertado, obviamente, si descubre que muchas teorías absolutamente nada atractivas pueden afirmar, todas ellas, que están tan bien apoyadas por las observaciones pasadas como la más atractiva. Pero ese problema no se me plantea a mí.

Por el contrario, si una teoría de gran capacidad explicativa le parece atractiva o prometedora a alguien, por la razón que sea (como le ocurrió a Galileo con su teoría no lunar), entonces tiene razón en seguir con ella y no perder la esperanza demasiado pronto, incluso ante dificultades internas e incluso ante aparentes refutaciones empíricas: al final puede conseguir formularla bien, a pesar de todo. Si no, cuanto mayor haya sido su esfuerzo intelectual por hacer frente a las críticas, más habrá aprendido de sus errores. Una cierta dosis de dogmatismo y de cabezonería es necesaria en la ciencia, si no queremos perder ideas brillantes que no sabemos cómo manejar o modificar en el primer momento.

Hay un lugar y una función, dentro de método crítico de la ciencia, incluso para los extremistas lunáticos. Escribí una vez que nuestras universidades no deberían tratar de crear eruditos y científicos, sino darse por satisfechas con un objetivo más modesto y más liberal: el de formar personas capaces de distinguir entre un charlatán y un erudito o un científico[50]. En seguida me corrigió L. E. J. Brouwer, que me dijo que ni siquiera esa fórmula era suficientemente liberal, puesto que podía interpretarse como un estímulo a esa nada liberal superioridad con la que el académico desdeña a menudo al de fuera. Indicó que había un lugar en la ciencia incluso para el charlatán y rechazó con toda razón todo lo que pudiera interpretarse como un apoyo a distinciones de esa índole.

Así, si fuéramos a dar razones metodológicas para condenar

[50] Véase la nota 6 del capítulo 11 de *La sociedad abierta (O. S.)*. (En esa nota —un ataque a la idea de producir expertos— se usa el término «experto» [expert] donde aquí hablo de «eruditos» [scholars]).

todas las teorías como la de que todos los cuervos son negros excepto los fotografiados durante un eclipse total de sol, podríamos fácilmente condenar una teoría de la mayor importancia. Podríamos dar razones *a priori* contra la ley de Eisntein de que hay grupos de estrellas fijas en la región del Zodíaco que siempre presentan las mismas distancias angulares relativas excepto si se las mide durante un eclipse total de sol, cuando se descubrirá que algunas de sus distancias angulares relativas son ligeramente diferentes. Si Galileo hubiera sabido los que es la fotografía, podría haber pensado fácilmente que esta ley de Einstein tenía un desagradable olor a astrología.

VI

Nuestra exposición ha sido, hasta el momento, lógica o metodológica o epistemológica. Puede resumirse diciendo que he sustituido el problema «¿cómo lo sabe usted? ¿cuál es la razón o la justificación de su aserción?» por el problema de «¿por qué prefiere usted esta conjetura a las conjeturas rivales? ¿cuál es la razón de su preferencia?».

Mientras que la respuesta al primer problema es «*no* sé», mi respuesta al segundo problema es, por regla general, que nuestra *preferencia* por una teoría mejor corroborada se defenderá racionalmente por medio de los argumentos que se han usado en nuestra discusión crítica, incluyendo, naturalmente, nuestra discusión de los resultados de las contrastaciones. Esos son los argumentos de los cuales el grado de corroboración trata de proporcionar un informe resumido.

De este modo se resuelve el problema lógico de la inducción.

Naturalmente, quedan muchos problemas que pueden considerarse aspectos o fases del problema de la inducción; y algunos de ellos pueden ser problemas lógicos. El principal problema que queda —yo lo llamo la cuarta fase del problema de la inducción— es, sin embargo, de un carácter diferente, a pesar de su estrecha conexión con los problemas ya discutidos.

5. POR QUE LA CUARTA ETAPA DEL PROBLEMA ES METAFISICA

La cuarta etapa del problema es metafísica. El desafío contenido en ella puede formularse como sigue:

Existen leyes naturales verdaderas. Esto lo sabemos, y lo sabemos por experiencia. Hume dice que no lo sabemos; pero a

pesar de lo que él dice, lo sabemos: nuestra creencia de que existen leyes naturales verdaderas está basada, sin duda, de un modo u otro, en las regularidades observadas: en el cambio del día y la noche, el cambio de las estaciones y en otras experiencias similares. Así pues, Hume tiene que estar equivocado. ¿Puede usted mostrar *por qué* está equivocado? Si no, no ha resuelto usted su problema.

Simpatizo de todo corazón con el espíritu de este desafío, aunque no estoy totalmente de acuerdo con su formulación.

Desde el principio debemos advertir que hay muchas maneras diferentes de interpretar la afirmación de que existen leyes naturales verdaderas. Por ejemplo:

1) Hay (en la actualidad) al menos un enunciado universal verdadero que describe las regularidades invariables de la naturaleza.

2) Algún posible enunciado universal que describe las regularidades invariables de la naturaleza (se haya expresado ya o no) es verdadero.

3) Existen regularidades en la naturaleza (tanto si se expresan jamás o si son expresables como si no).

Aunque todas estas afirmaciones están *relacionadas* con el problema de la inducción de Hume, las cuestiones ligadas a ellas son, sin embargo, bastante diferentes, y en varios aspectos diferentes.

Primero, lo que Hume mostró fue que no podíamos derivar una ley universal como «todos los cisnes son blancos» por muchas que fueran las observaciones realizadas (o los enunciados de observaciones) de cisnes blancos. Pero los enunciados 1-3) que acabamos de enumerar no son universales: son enunciados existenciales singulares. De modo que se plantea la cuestión de cómo tal enunciado existencial singular se relaciona deductivamente con la observación o quizá con las reflexiones sobre nuestra experiencia.

Nuestra aserción —«Existen leyes naturales verdaderas»— al ser existencial, no se refiere siquiera a ninguna ley física determinada, sino que solamente afirma que al menos una ley tal es verdadera. Esto tiene una importancia considerable. Yo, por lo menos, no estaría dispuesto a señalar a ninguna ley determinada de la física y decir: «Esta ley es verdadera, en su formulación y su interpretación actuales: Estoy seguro de que nunca será falsada ni modificada ni reconocida como sólo condicionalmente válida, o como válida sólo dentro de ciertos límites». Al mismo tiempo, creo que al menos *alguna* de las leyes de nuestro actual sistema de física es verdadera en ese sentido; yo diría incluso que muchas lo son, si incluimos las que tienen un bajo nivel de universalidad.

Segundo, nuestra aserción, al menos en sus dos primeras interpretaciones, no pertenece a la física. Más bien habla *de* las teorías de la física (o quizá de la ciencia en general). Esto puede expresarse diciendo que pertenece a la meta-teoría de la física. (Pertenece a lo que Tarski llama la «semántica» de la ciencia física.) Mientras que los enunciados de la ciencia son sobre objetos no lingüísticos, nuestra aserción es, pues, sobre objetos lingüísticos. De forma que pertenece a un lenguaje (un «meta-lenguaje») en el cual podemos hablar sobre algún otro lenguaje (el «lenguaje-objeto»), que, a su vez se refiere al mundo. El enunciado «existe alguna ley natural verdadera» es una conjetura sobre el mundo y también un comentario sobre las leyes naturales. Mientras que el problema original de Hume se refiere a la relación lógica entre una ley natural y alguna experiencia observacional, nuestro nuevo problema se refiere a la relación, lógica o de otra clase, entre los comentarios sobre las leyes naturales y los comentarios (o reflexiones) sobre experiencias observacionales.

Podría objetarse aquí que nuestra aserción de que hay al menos una ley verdadera de la naturaleza pertenece, a pesar de lo que acabo de decir, a la ciencia. El argumento podría ser que las leyes de la naturaleza pertenecen a la ciencia y si *a* es una ley de la naturaleza, entonces «*a* es verdadera» se deduce de «*a*» (según la definición de Tarski de «verdadero») y a partir de «*a* es verdadera» y «*a* es una ley de la naturaleza», obtenemos, desde luego, «existe una ley verdadera de la naturaleza». Admito la corrección de la educación; «existe una ley verdadera de la naturaleza» se sigue, en efecto, de cualquier ley científica. Sin embargo, como todas las leyes científicas son conjeturales, el enunciado «existe una ley verdadera de la naturaleza» no tiene por qué ser verdadero a su vez; hereda el carácter conjetural de la ley *a*. Al mismo tiempo, *no* adquiere carácter científico simplemente por seguirse de una ley científica: enunciados incontrastables y, por tanto, no científicos, se siguen, de modo trivial, de cualquier enunciado contrastable. Y, en especial, el enunciado «existe una ley universal científica (o contrastable) verdadera» es, a su vez, incontrastable. Desde luego, se sigue de cualquier aserción (conjetural) de una ley científica, como acabamos de ver.

Otro problema que también debilita el interés de esta derivación es que los argumentos usados en nuestra discusión científica para la valoración de conjeturas científicas no pueden usarse en apoyo de la conjetura de que «existe una ley verdadera de la naturaleza». Porque esos argumentos apoyan simplemente nuestra *preferencia* por una u otra ley y no establecen, ni apoyan, la concepción de que alguna de ellas sea *verdadera.*

Tercero, en su tercera interpretación (en la que estoy especialmente interesado aquí), nuestra aserción toma un carácter *metafísico,* en varios de los muchos sentidos usuales del término «metafísico» y en un sentido en el que puede usarse en oposición a «lógico», «metodológico» o «epistemológico».

A diferencia de las aserciones puramente metodológicas y también *a diferencia* de las aserciones puramente metalingüísticas, pero *al igual* que las conjeturas de la propia ciencia, nuestra aserción puede interpretarse como una conjetura *sobre el mundo.* Puede interpretarse que la afirmación de que existe una ley verdadera de la naturaleza significa que el mundo no es completamente caótico, sino que tiene ciertas regularidades estructurales «incorporadas», por decirlo así. De ahí que pertenezca a una teoría de la estructura del mundo, a una especie de cosmología general: es una conjetura de una cosmología metafísica.

Está claro que nuestra aserción, al ser existencial, no puede ser contrastada empíricamente; no es falsable; y tampoco es verificable, puesto que ninguna ley es verificable. Como nuestra aserción es irrefutable, podemos describirla con certeza como «metafísica» en el sentido técnico que se le da a este término en *L. I. C.* (Compárese con la sección 6.)[51] Y al no ser ni falsable ni verificable es, presumiblemente, «metafísica» también en el sentido de los positivistas.

Es metafísica también en el sentido tradicional de la palabra, puesto que trata de cuestiones que se consideran como características de la metafísica. Trata de la misma clase de cuestión que, por ejemplo, *el principio de causación universal,* una de cuyas posibles formulaciones es: «Para todo suceso de este mundo, existen leyes universales verdaderas y condiciones iniciales verdaderas a partir de las cuales pueden deducirse un enunciado que describa el suceso en cuestión». También esto es una aserción sobre el mundo y su estructura.

Puede objetarse que cualquiera que sea la interpretación metafísica posible de nuestra aserción, sigue perteneciendo principalmente a la metodología o a la teoría del conocimiento. Puede sostenerse que para buscar leyes verdaderas (como hacemos) debemos *presuponer* en nuestra búsqueda la existencia de dichas leyes. Así, «existen leyes verdaderas de la naturaleza» es una presuposición metodológica.

Pero esta objeción no es concluyente. Porque uno puede

[51] Véase *L. I. C., passim; El desarrollo del conocimiento científico (C. & R.),* esp. capítulo 11 y Apéndice al capítulo 10. Véase también mi artículo «Indeterminism in Quantum Physics and in Classical Physics», en *The British Journal for the Philosophy of Science* 1, 1950, pp. 117 ss.

muy bien buscar algo que no existe, y sin suponer, presuponer ni posturlar su existencia. Por ejemplo, cuando contrastamos una ley, buscamos un contra-ejemplo de ella. Pero ni suponemos, ni presuponemos ni postulamos la existencia de tal contra-ejemplo. De hecho, puede no existir tal contra-ejemplo: la ley que estamos contrastando puede ser verdadera.

Entonces, aunque no presupongamos, o supongamos que «existen leyes verdaderas de la naturaleza» podemos —y efectivamente, sin duda, lo hacemos— creerlo. Y quizá esta creencia sea psicológicamente importante en nuestra búsqueda de leyes verdaderas. Pero incluso esto no hace de ella un presuposición *metodológica*; hace, simplemente, una presuposición psicológica.

Por cierto, yo comparto esta creencia y la considero más razonable que ninguna de las alternativas que conozco. La mejor manera de entender —y evaluar— esta creencia es considerarla como una conjetura metafísica sobre la estructura del mundo.

Antes de ocuparnos de esta cuestión metafísica, sin embargo, hay otra objeción más o menos metodológica; una que también explota la conexión, mencionada más arriba, con la ley de la causación universal.

Muchos filósofos han sostenido que el problema de la verdad de la ley de causación universal (o de la «Uniformidad de la Naturaleza» que es, tal vez, una fórmula más vaga aún) es equivalente al problema de Hume. Es decir, puede mantenerse que podemos usar la ley de causación universal como un principio de inducción cuya validez haría válidas las inferencias inductivas.

Esta sugerencia es, sin embargo, totalmente errónea. Quizá fuera excusable en los tiempos anteriores a Einstein, pero no lo es, en absoluto, después de él. Después de Einstein, es evidente que un principio inductivo —un principio que hiciera válida la inferencia inductiva— no puede existir. Porque si puede descubrirse que una teoría tan bien confirmada como la de Newton es falsa, entonces, obviamente, ni siquiera la mejor evidencia inductiva puede garantizar nunca la verdad de una teoría. En consecuencia, ningún principio inductivo que nos permita extraer inferencias inductivas será válido: será refutado por la primera refutación de una teoría que se haya inducido de acuerdo con el principio inductivo en cuestión.

Pero si no puede existir una solución positiva al problema de Hume en forma de un principio válido de inducción, entonces la ley de la causación universal —*sea ésta lo que fuere*— no puede ser un principio válido de inducción. El mismo resultado puede argüirse más directamente: la ley de la causación universal podría ser verdad y, no obstante, podríamos no hacer

ningún progreso científico, quizá porque las condiciones iniciales varían tan radicalmente que prácticamente no se repiten nunca ni siquiera aproximadamente, o a causa de la complejidad de las leyes o por otras razones. Así, la ley de la causación universal, incluso si fuera verdadera, carecería de significación metodológica. La regla metodológica, significativa e importante «buscad leyes naturales», no se sigue a ella. Tampoco se les promete éxito a aquellos que actúan de acuerdo con este imperativo.

Si la ley de la causación universal no tiene significación metodológica, no es sorprendente que carezca también de significación metodológica directa una aserción mucho más débil, de lo que nos ocupamos aquí, «existen regularidades en la naturaleza» —la aserción cuya validez se cuestiona en la cuarta etapa de nuestro problema—. Incluso si supiéramos con certeza que hay regularidades en la naturaleza, los argumentos de Hume contra la inducción serían válidos. Millones de observaciones de hombres que hablan inglés no establecen que todos los hombres hablen inglés: ninguna sucesión de observaciones de los elementos de una muestra pueden decirnos que nos encontramos antre una muestra *adecuada*. Por otro lado, incluso si supiéramos que no había regularidades invariables —incluso si existieran contra-ejemplos para *todas* las leyes aparentes— seguiría teniendo mucho sentido *intentar* racionalizar un mundo tan fundamentalmente irracional, en la medida de lo posible, por el método crítico de prueba y error.

Así pues, nuestro problema metafísico es en gran medida académico y tiene un carácter bastante diferente del de los problemas lógicos y metodológicos que hemos resuelto. El hecho de que la cuarta etapa, o etapa metafísica, de nuestro problema continúe sin resolverse es, por consiguiente, perfectamente compatible con mi afirmación de que he resuelto enteramente el problema de Hume de la inducción sobre una base lógica, metodológica y epistemológica. La solución de la cuarta fase no es necesaria para establecer mi afirmación.

El argumento lógico de Hume contra la inducción simplemente no atañe directamente a nuestra aserción metafísica de que existen regularidades en la naturaleza. No obstante, es perfectamente cierto que tendremos que defender esta aserción metafísica contra Hume, pero no contra su *lógica;* más bien, contra su *metafísica.*

Estamos ahora en una posición que nos permite reformular con más claridad la cuarta fase, o fase metafísica, del problema de la inducción. Podemos hacerlo resaltando que es un aspecto del *problema de mañana.*

He convenido (pueden decirme) en que la conjetura meta-

física de que hay leyes naturales verdaderas, en el sentido de «existen regularidades», es mejor que las alternativas que se le conocen y, por tanto, que es razonable creer en ella. Pero si hay leyes naturales, entonces puede argüirse como sigue: si *a* es una de tales leyes, *a* continuará siendo aplicable, u operando, en el futuro, digamos que *mañana*. Pero, ¿cómo puedo decir que es razonable creer esto si estoy de acuerdo con Hume? Hay que admitir que una conjetura científica puede ser más razonable que otra, a la luz del estado actual de nuestra discusión crítica; pero, si admitimos que Hume tiene razón, ¿por qué vamos a aceptar que es más razonable creer en esta conjetura metafísica que en sus alternativas? ¿Por qué no podrían cambiar lentamente, por ejemplo, *todas* las regularidades aparentes? Mi argumento anterior no parece aplicable aquí, porque era *metodológico:* mostraba por qué, si deseamos explicar el mundo con los métodos de la ciencia, cualquier cambio en las regularidades conocidas tendría que ser explicado con ayuda de nuevas leyes (conjeturadas). Pero ahora nos enfrentamos con cuestiones diferentes. Las cuestiones que nos planteamos ahora son: ¿por qué no puede ocurrir que fracasen totalmente la ciencia y su método mañana porque fallen *todas* las regularidades, tanto si han sido tomadas en consideración previamente como si no? ¿Y por qué es razonable creer que esto no va a ocurrir, y que, incluso aunque tengan lugar tales cambios mañana, que fracasen la ciencia y su método, habrá, sin que nosotros lo sepamos, al menos una regularidad que continuará operado porque sea verdaderamente invariable?

La discusión del problema metafísico nos ocupará durante algún tiempo. Pero como primer paso hacia la solución, quiero señalar lo ingenua que es cualquier formulación del problema que use términos temporales tales como «mañana» o «el futuro». Porque cualquiera de esas formulaciones da ingenuamente por sentada una regularidad: *el orden del tiempo.*

De hecho, todas las formulaciones «¿se asemejará el futuro al pasado?» se fundan ingenuamente en la aceptación acrítica e inconsciente de una *teoría del tiempo* intuitivamente «natural» y sin embargo sumamente sospechosa. Es una teoría que fue sostenida y expresada, por ejemplo, por San Agustín y que Newton fue uno de los primeros en formular explícitamente con las palabras: «El tiempo absoluto, verdadero y matemático fluye igualmente de sí mismo y por su propia naturaleza, sin relación con ninguna cosa externa»[52].

Los que formulan el problema de la inducción en términos temporales (como «el futuro» o «mañana») presuponen ins-

[52] Isaac Newton, *Philosophiae Naturalis Principia mathematica*, 1687.

conscientemente esta teoría del tiempo u otra similar, sin ser conscientes, al parecer, de su carácter problemático. Porque reflexionan sobre si las leyes de la naturaleza pueden cambiar o no, en el sentido de que en el futuro, o mañana, haya regularidades diferentes de las que han estado en vigor hasta ahora. Pero eso entraña la teoría de que *el futuro, o mañana, llegará independientemente de un cambio en las leyes de la naturaleza.* Así pues, suponen un flujo del tiempo que es independiente de todo lo que pase e independiente de las leyes de la naturaleza. Suponen, por tanto, precisamente, lo que Newton trata de describir. Pero son más ingenuos que Newton y no advierten que, en la formulación misma de su problema, suponen inconscientemente que ciertas leyes de la naturaleza —las leyes del flujo del tiempo absoluto— están exentas de la duda de Hume. En otras palabras, aunque creen ser empiristas e inductivistas, suponen, con Kant, que esas leyes del tiempo son válidas *a priori*.

Esta concepción es equivalente a una cosmología metafísica en las líneas siguientes: existe el tiempo (y, presumiblemente, el espacio) y todas las cosas en la naturaleza ocurren *en* el tiempo. El mundo es una totalidad de sucesos, no de cosas (fue Heráclito el primero en ver esto[53]) y los sucesos son esencialmente *en* el tiempo (y, presumiblemente, *en* el espacio).

Esta cosmología fue atacada por Leibniz y Berkeley que, por separado, propusieron *Una teoría relacional del tiempo y el espacio.* Se consideró el tiempo como un sistema para ordenar relaciones (tales como *antes, después, simultáneamente*) que se daban entre *sucesos*; y el espacio como un sistema para ordenar relaciones que se daban entre cosas. El mundo es aquí, nuevamente, una totalidad de sucesos. Pero estos sucesos no son en un tiempo cuya existencia es una condición para que exista un suceso. Antes bien, sólo existe la totalidad de los sucesos, junto con sus relaciones temporales; y «*tiempo*» es simplemente una palabra, un nombre para el sistema abstracto de esas relaciones temporales. Esta concepción más compleja puede, quizá, no ser correcta, pero se ha aceptado generalmente en ciencia, incluso antes de Einstein.

6. EL PROBLEMA METAFISICO

El principio metafísico de la causación universal es uno en el que yo no creo. (*A fortiori,* yo no creo en el principio más

[53] Una defensa de esta interpretación algo polémica se encontrará en mi *Sociedad abierta (O. S.),* nota 2 al capítulo 2, y en *El desarrollo del conocimiento científico (C. & R.),* cap. 5 (y el Apéndice al cap. 5).

fuerte del determinismo «científico» que se discutirá en el tomo II de este *Post Scriptum.*) Pero sí que creo, como ya he indicado, en el principio, mucho más débil, «existe al menos una ley verdadera de la naturaleza». Delinearé una variedad de argumentos en su favor. Después llamaré la atención sobre algunas de las dificultades de la posición metafísica que propongo que se adopte.

Reconsideremos, en primer lugar, mi respuesta a la primera etapa del problema, dada en las secciones 2 y 4. Allí hice hincapié en que las teorías científicas son suposiciones o conjeturas *que pueden o no ser verdaderas* y nunca podemos saber que una teoría es verdadera, aunque lo sea. Lo que quiero resaltar ahora es esto: el hecho de que no sepamos y no podamos saber que una teoría es verdadera no es, en sí mismo, una razón para que la teoría no sea verdadera. Puede ser una razón para dejar en suspenso la creencia, pero no es ciertamente, una razón para la incredulidad; es decir, para creer que la teoría es falsa[54].

Podemos reconsiderar ahora la respuesta a la tercera etapa, dada en la sección 4. Dije allí que es razonable actuar sobre la base de una teoría científica discutida minuciosamente y bien contrastada (y creer en ella) siempre que estemos dispuestos a cambiar de opinión a la luz de nuevos argumentos; por ejemplo, de nueva evidencia empírica.

Hasta cierto punto, esta observación resuelve también la cuarta fase del problema. Porque creer en un enunciado y creer en la verdad de un enunciado es lo mismo. (Esto concuerda con la teoría de Tarski.) Puede, por tanto, considerarse razonable creer que existe una ley verdadera de la naturaleza, siempre que exista una ley de la naturaleza discutida minuciosamente y bien contrastada. Como, de hecho, tenemos un número considerable de leyes de la naturaleza que han sido minuciosamente discutidas y bien contrastadas, hay, efectivamente, razones empíricas para la creencia de que existe al menos una ley verdadera de la naturaleza.

Puede uno tener la sensación, sin embargo, de que esta respuesta no es enteramente satisfactoria. Y el mencionar a Hume en nuestra formulación original del problema contiene quizá una clave de la razón de esta insatisfacción: Hume, si no un idealista declarado, era por lo menos un escéptico respecto a la realidad del mundo físico. Su escepticismo estaba estrecha-

[54] En Hume, el conocimiento es un tipo de creencia verdadera justificada. Todo este enfoque choca con el mío. Si hablo aquí de «creencia» es en un sentido diferente, el sentido, más bien, de mi *Conocimiento Objetivo (Obj. Kn.).* Para mí, el conocimiento —es decir, el conocimiento conjetural— es objetivo: está fuera, es un producto de nuestras mentes y no un estado de nuestras mentes. No me tomo en serio el «problema» de la creencia.

mente relacionado con sus concepciones sobre la inducción. Admitía la fuerza de nuestra creencia en un mundo físico ordenado por leyes, pero afirmaba que esta creencia era infundada. Esto sugiere que la cuarta etapa de la cuestión debería haber sido: «Creo que vivimos en un mundo real, un mundo que presenta cierto tipo de orden estructural que se nos aparece bajo la forma de leyes. ¿Puede usted mostrarme que esta creencia es razonable?».

La cuestión que se plantea aquí es la del realismo metafísico, en una forma que no resalta tanto la existencia de cuerpos físicos como la existencia de leyes. Porque los cuerpos físicos son sólo un aspecto de la estructura legiforme del mundo, que garantiza, sola, su permanencia (relativa); lo que significa, por otro lado, que la existencia de cuerpos físicos (sobre la que Hume es tan escéptico) entraña la de regularidades físicas objetivas (*cf.* sección 16).

7. REALISMO METAFISICO

> *Un desastroso miedo a la metafísica... [es la] enfermedad de la filosofía empírica contemporánea... Este miedo parece ser el motivo de interpretar, por ejemplo, una «cosa» como un «manojo de cualidades», «cualidades» que pueden descubrirse, se supone, entre la materia prima de nuestros sentidos... Yo, por el contrario, no creo que esté implicado ningún tipo peligroso de metafísica en admitir la idea de una cosa física (o un objeto físico) como una noción autónoma en el sistema, junto con la estructura espacio-temporal adecuada a ella.*
>
> ALBERT EINSTEIN

Afortunadamente, o quizá desgraciadamente, *L. I. C.* no era un libro sobre metafísica, al menos sobre metafísica del «tipo peligroso» al que se refiere Einstein[55]. Tampoco lo es este *Post Scriptum*. Sin embargo, en *L. I. C.* declaré que creía en el realismo metafísico (*cf.* el segundo párrafo de la sección 79, y el final de las secciones 4 y 28). Y sigo creyendo en el realismo metafísico.

---

[55] Al menos no metafísica del «tipo peligroso» al que se refiere Einstein en el lema, traducido de su contribución a P. A. Schilpp, ed., *The Philosophy of Bertrand Russel*, 1944, p. 230.

En ningún lugar se usa el realismo metafísico en apoyo de ninguna de las soluciones propuestas en *L. I. C.* (En esto, mi método difiere de la práctica habitual de los idealistas que, desde Berkeley y Hume hasta, digamos, Reinchenbach usan sus concepciones metafísicas en apoyo de sus teorías epistemológicas.) No es una de las tesis de *L. I. C.*, ni tampoco tiene el papel de una presuposición en ningún lugar. Y sin embargo, está ahí. Forma una especie de trasfondo que da sentido a nuestra búsqueda de la verdad. La discusión racional, es decir, la argumentación crítica con el propósito de acercarse más a la verdad, carecería de sentido sin una realidad objetiva, un mundo de cuyo descubrimiento hacemos nuestra tarea: desconocido, o en gran medida desconocido: un desafío a nuestro ingenio, valentía e integridad intelectuales. No hay en *L. I. C.* compromiso con el idealismo, ni siquiera con la concepción de que conocemos el mundo solamente a través de nuestras observaciones, la concepción que tan fácilmente conduce a la doctrina de que todo lo que conocemos, o podemos conocer, son nuestras propias experiencias observacionales (*cf. L. I. C.*, capítulo V).

Este sólido, aunque esencialmente implícito, realismo que impregna *L. I. C.* es uno de sus aspectos por los que siento cierto orgullo. Es también uno de los aspectos que la vincula a este *Post Scriptum*, cada uno de cuyos tomos ataca a uno y otro de los enfoques subjetivistas o idealistas del conocimiento. Por tanto, puede no estar fuera de lugar discutir aquí y en las próximas nueve secciones, aunque sea a grandes rasgos, algunos problemas metafísicos como tales, sobre todo en vista de que están relacionados, en diversas formas, con el problema de la estructura y de la condición de la ciencia (o del «conocimiento científico» en el sentido explicado en la sección 1). Esta discusión va a ocuparnos hasta la sección 16.

La intención de los filósofos empíricos, desde Bacon hasta Hume, Mill y Russell, era práctica y realista. Con la posible excepción de Berkeley, todos querían ser realistas con los pies en la tierra. Pero sus epistemologías subjetivistas chocaban con sus intenciones realistas. En lugar de atribuir a la experiencia de los sentidos la importante, pero limitada capacidad de contrastar, o comprobar, nuestras teorías sobre el mundo, esas epistemolgías sostenían «la teoría de que todo conocimiento se deriva de la experiencia de los sentidos»[56]. E identificaban «se deriva» o bien con «se deriva inductivamente» o, incluso con

[56] Esta es una definición del empirismo, de la *Encyclopaedia Britannica*, citada por Russell al principio de su artículo titulado «The Limits of Empiricism»; véase la nota 1 a la sección 1.

mayor frecuencia, con «se origina». Nunca vieron claramente que no es el origen de las ideas lo que debe interesar a los epistemólogos, sino la verdad de las teorías; y que el problema de la verdad o falsedad de una teoría sólo puede surgir, obviamente, *después* de que la teoría nos haya sido presentada —es decir, *después* de que alguien le haya dado origen, de una u otra forma— y que la historia de su origen apenas tiene nada que ver con la cuestión de su verdad. (Recuerdo perfectamente a un viejo campesino en las montañas del Tirol que daba por sentado que el trueno era el ruido producido por la colisión de nubes pesadas y el relámpago, una chispa muy caliente debida a su fricción. Tengo pocas dudas sobre el hecho de que el *origen* histórico de esta teoría tan directa debe ser menos sospechoso —es decir, está más cerca del modelo inductivo— que el de la teoría más compleja que han adoptado los modernos meteorólogos.)

La creencia del filósofo empirista de que «todo conocimiento se deriva de la experiencia de los sentidos» lleva necesariamente a la concepción de que todo conocimiento tiene que ser conocimiento de nuestra experiencia sensorial actual (las «ideas de las impresiones» de Hume) o de nuestra pasada experiencia sensorial (las «ideas de reflexión» de Hume). Así, todo conocimiento se convierte en conocimiento de lo que está ocurriendo en nuestras mentes. *Sobre esta base subjetiva, no puede construirse ninguna teoría objetiva:* el mundo viene a ser la totalidad de mis ideas, de mis sueños.

La doctrina de que el mundo es un sueño mío —es decir, la doctrina del idealismo— es irrefutable. Puede abordar toda refutación interpretándola como un sueño (del mismo modo que el psicoanálisis puede abordar toda crítica psicoanalizándola). Pero la creencia, muy extendida, de que la irrefutabilidad de una teoría es un argumento en su favor, es errónea. *La irrefutabilidad no es una virtud sino un vicio.* Esto es válido también para el realismo, desgraciadamente: porque también el realismo es irrefutable. (La refutación del realismo es sólo un sueño de un idealista. La muerte, sueña él, puede ser el despertar que nos demuestre finalmente que mientras vivíamos sólo estábamos soñando. Pero como argumento, esto ni siquiera tendería a refutar el realismo: si al despertar fuésemos a percatarnos de que sólo habíamos estado soñando, ello ocurriría porque seríamos capaces de distinguir entre el sueño y la realidad. Pero esto es precisamente lo que el idealista dice que no podemos hacer.) De la irrefutabilidad del idealismo se sigue la indemostrabilidad del realismo y viceversa. Ambas teorías son indemostrables (y, por tanto, sintéticas) y también irrefutables: son «metafísicas».

Pero hay una diferencia importantísima entre ellas. El idealismo metafísico es falso y el realismo metafísico es verdadero. No «sabemos» esto, naturalmente, en el sentido en que podemos saber que 2 + 3 = 5; es decir, no lo sabemos en el sentido del conocimiento demostrable. Tampoco lo sabemos en el sentido del «conocimiento científico» contrastable. Pero esto no significa que nuestro conocimiento no sea razonado o razonable. Por el contrario, no hay conocimiento factual que esté apoyado por más argumentos o por argumentos más fuertes (aunque no sean concluyentes).

Antes de entrar de lleno a considerar los argumentos positivos en apoyo del realismo metafísico, discutiré algunos argumentos negativos: éstos apoyan al realismo por medio de la crítica del idealismo.

Desde el punto de vista de una epistemología subjetiva o idealista, la forma más extrema del idealismo es el solipsismo. El argumento epistemológico en favor del idealismo es que todo lo que yo conozco son mis propias experiencias, mis propias ideas. Sobre otras mentes no puedo conocer nada directamente. De hecho, mi conocimiento sobre otras mentes tendría que depender de mi conocimiento sobre cuerpos: no tenemos conocimiento empírico de espíritus incorpóreos. Si los cuerpos no son más que partes de mi sueño, otras mentes deben serlo más aún.

El problema de las otras mentes ha sido discutido interminablemente en años recientes, en gran medida en términos epistemológicos. Confieso que no he leído todas esas discusiones y no es, pues, imposible que mi sencillo argumento en favor de la existencia de otras mentes haya sido utilizado anteriormente por otros (aunque no creo que haya sido así). Me satisface por completo este argumento, quizá porque siempre recuerdo que, en este tipo de investigación, no hay argumentos concluyentes.

Mi argumento es éste: yo sé que yo no he creado la música de Bach ni la de Mozart; que no he creado los cuadros de Rembrandt ni los de Boticelli. Estoy totalmente seguro de que nunca podría hacer nada así: simplemente, no lo tengo en mí. (Esto lo sé especialmente bien, puesto que hice muchos intentos de copiar a Bach; eso me hizo apreciar más su capacidad de invención.) Sé que no tengo la imaginación necesaria para escribir nada como *La Ilíada*, o *El Infierno* o *La Tempestad*. Si es posible, soy aún más incapaz de dibujar un tira cómica corriente, de inventar un anuncio para la televisión o de escribir algunos de los libros sobre la justificación de la inducción que me veo obligado a leer. Pero según la hipótesis solipsista, todas esas creaciones serían las de mis propios sueños. Serían criaturas de mi propia imaginación. Porque no habría otras

mentes: no habría nada sino mi mente. Sé que esto no puede ser verdad.

El argumento no es concluyente, naturalmente. Quizá yo me minusvalore a mí mismo (y al mismo tiempo, me sobrevalore) en mi sueño. O puede no ser aplicable a la categoría de la creación. Todo esto está entendido. No obstante, el argumento me satisface plenamente.

Sin duda, mi argumento es un poco similar al de Descartes de que una mente finita e imperfecta no puede crear por sí misma la idea de Dios, pero encuentro mi propio argumento más convincente. La analogía con el argumento de Descartes sugiere, sin embargo, una simple extensión al mundo físico. Sé que soy incapaz de crear, con mi propia imaginación, nada tan bello como las montañas y los glaciares de Suiza, y ni siquiera algunas de las flores y los árboles de mi propio jardín. Sé que el nuestro es un mundo que yo nunca creé.

Sólo puedo repetir que este argumento me satisface; quizá porque nunca lo he necesitado realmente: no pretendo hacer creer que yo he dudado alguna vez de la realidad de otras mentes o de los cuerpos físicos. En realidad, cuando pienso en este argumento, no puedo dejar de sentir que el solipsismo (o, más en general, la duda en la existencia de otras mentes) no es tanto una forma de epistemología como de megalomanía.

Y nada más sobre el solipsismo y sobre otras formas de epistemología idealista que ponen en duda la existencia de «otras mentes». Me parece muy posible que argumentos como el que acabo de proponer impidieran que Berkeley se convirtiera en un solipsista: como era cristiano, sabía que él no era Dios. De modo que llegó a la concepción de que existían otras mentes además de la suya y que era Dios quien nos hacía percibir esa cosa tan llena de esplendores: el mundo de nuestra experiencia.

La versión de Berkeley del idealismo es tan irrefutable como cualquier otra y tiene poco de recomendable. Incluso suponiendo que el argumento epistemológico favorezca la tesis solipsista, está claro que un recurso al argumento epistemológico no puede ser convincente una vez que se admiten realidades que no son percibidas, tales como Dios y otras mentes humanas. El intento de Berkeley de conciliar el idealismo epistemológico y el cristianismo lleva a un aparente compromiso que, de hecho, perjudica a ambos. (El cristianismo resulta perjudicado porque el sufrimiento físico de Cristo no le es infligido ya por los hombres, sino por la acción inmediata de la deidad)[57].

[57] Véase también la sección 11, más abajo, el texto correspondiente a la nota 95.

Ninguno de estos argumentos debería ser necesario. El realismo es tan obviamente verdadero que incluso un argumento directo, como el que se ha presentado aquí, es de cierto mal gusto. Tiene algo de trillado y de rancio que me recuerda un poco a una costumbre que detesto: la de filosofar sin un problema real. «Se ha dicho que todos los hombres inteligentes pertenecen a la misma religión y que ningún hombre inteligente dice jamás cuál es esa religión»[58]. Tengo la impresión de que continuar con mi argumento sería olvidar la segunda mitad de esta máxima tan sabia.

Sería injusto, hay que admitirlo, decir que los idealistas no tenían de hecho ningún problema real. Su problema era la «justificación» (positiva) de nuestro conocimiento y cayeron en una trampa: en su propio descubrimiento de que era imposible «justificar» el realismo. W. W. Bartley III me ha señalado que es injusto juzgarles desde un punto de vista que descarta la totalidad del programa de la «justificación» positiva como fútil y lo sustituye por un programa de crítica (*cf.* sección 2, más arriba). Acepto esta advertencia. Pero ¿cuál, entre los filósofos idealistas, ha hecho alguna vez hincapié en la cuestión de que, incluso si el realismo es verdadero, no podemos justificarlo en el sentido de ellos, del mismo modo que no podemos justificar el idealismo si el realismo es falso? ¿Y que, por consiguiente, la imposibilidad de «justificar» el realismo no habla en contra de su verdad? Y ¿cuál de ellos dejó claro que, puesto que esta situación caracteriza a la estructura lógica del probblema, es obviamente bastante fútil usar como argumento en contra del realismo el de que no puede ser «justificado» o, en realidad, cualquier argumento similar?

La exasperante ranciedad de los argumentos de idealistas y sensacionalistas proviene de que no ven las limitaciones lógicas inherentes a sus programas justificacionistas. No ven, sencillamente, que incluso una prueba lógica de la imposibilidad de justificar el realismo no constituiría una justificación de su negación[59].

Mis argumentos son aplicables no sólo al solipsismo y al idealismo berkeleyano, sino a *todas* las otras formas de esta enfermedad (en la medida en que las conozco), especialmente a las diversas formas de positivismo y fenomenalismo y también

---

[58] *The Note-Books of Samuel Butler* (Shrewsbury Edition, 1926), p. 229.

[59] Bartley me ha llamado la atención sobre el hecho de que una cuestión similar fue planteada por Ralph Barton Perry en «The Ego-Centric Predicament», *Journal of Philosophy, Psychology and Scientific Methods* 7 (1910), pp. 5-14. El argumento de Perry era que si suponemos, o admitimos, el hecho del procedimiento ego-céntrico, entonces *nada se sigue* de éste respecto a la verdad o falsedad del realismo o del idealismo.

al llamado «monismo neutral» de Willian James, Ernst Mach y Bertrand Russell, como mostraré en la próxima sección.

Los pocos deseos que tenía Bertrand Russell de ser idealista pueden verse en el bello pasaje en el que describe sus sentimientos después de su conversión al realismo: «pensábamos que es real *todo* lo que el sentido común, no influenciado por la filosofía o la teología, supone real. Con la sensación de escapar de una prisión, nos permitíamos pensar que la hierba es verde, y que el sol y las estrellas existirían aunque nadie fuera consciente de su existencia...»[60].

Pero, como creía en la inducción, Russell encontró que su epistemología no cumplía lo prometido. «La teoría del conocimiento», nos dice, «posee una cierta subjetividad esencial: pregunta '¿cómo se *yo* lo que sé?' y parte, inevitablemente, de la experiencia personal. Sus datos son egocéntricos, lo mismo que las primeras fases de su argumentación. Hasta el momento, no he pasado de las primeras fases y, por tanto, parece que tenga una visión más subjetiva que la que tengo realmente»[61]. Este pasaje es interesante a causa de la franqueza con la que revela que no ha alcanzado aún la meta realista y también por la claridad con la que localiza el error fundamental: si admitimos que nuestro conocimiento es conjetura, entonces, la pregunta fundamental de Russell: «¿cómo sé yo lo que sé?» resulta mal formulada, porque esta pregunta, en términos de conocimiento, equivale a preguntar: «¿Ha dejado usted ya de pegar a su mujer?». Supone que yo sé y, por consiguiente, que la inducción es válida. La pregunta aparentemente análoga, en términos de conjetura, tal como: «¿cómo (o por qué) conjeturo yo lo que conjeturo?» no es en absoluto análoga: esta pregunta es psicológica; no tiene efecto epistemológico. Así pues, la respuesta adecuada a la pregunta de Russell es: «*no* sé; y respecto a las conjeturas, no importa cómo o por qué conjeturo lo que conjeturo. No estoy tratando de demostrar que mis conjeturas son correctas, sino que estoy deseando que sean criticadas para sustituirlas por otras conjeturas mejores, si es posible. Y si *ustedes* tienen tantas dudas respecto a mis conjeturas como tengo yo, espero que me ayuden criticándolas sin piedad»[62].

En el momento en que sustituimos la idea de conocimiento por la idea de conjetura, la aparentemente «esencial subjetividad» de la teoría del conocimiento desaparece. Quizá algún conocimiento (*epistèmè*?) tenga que ser explicado, esencialmente, sobre una base subjetiva, sobre la base de lo que *yo* sé

[60] P. A. Schilpp, ed., *The Philosophy of Bertrand Russell*, 1944, p. 12.
[61] *Op. cit.*, p. 16.
[62] Véase también el final de la sección 4, más arriba.

con seguridad. Pero las conjeturas, por el contrario, son proposiciones y, como tales, pueden enfrentarse a las contraproposiciones de cualquier otro. El problema de su base subjetiva en nuestros sentidos («no hay nada en mi mente que no estuviera antes en mis sentidos») no necesita plantearse. *Nos movemos, desde el principio mismo, en el terreno de la intersubjetividad, del toma y daca de proposiciones y de críticas racionales.*

El problema fundamental de Russell necesita, pues, ser reformulado en términos de conjeturas; en términos del carácter hipotético del conocimiento que (en otro contexto) él sería el primero en resaltar. Por tanto, creo que yo tenía razón al presentar este argumento como central cuando repliqué al artículo de Russell en la Sociedad Aristotélica, ocho años antes de que él publicase los pasajes citados aquí (*cf.* sección 1).

A continuación del pasaje ya citado, Russell, escribe: «Si alguna vez tengo tiempo para emprender otra investigación seria de un problema filosófico, intentaré analizar las inferencias de la experiencia al mundo de la física, *suponiéndolas capaces de tener validez* y tratando de descubrir qué principios de inferencia, si son verdaderos, las harían válidas»[63]. Así, Russell estaba dispuesto a adoptar lo que Kant llamó un método «transcendental»: el método de tomar el conocimiento científico como un hecho y buscar los principios que explicasen cómo era posible ese hecho. El resultado (presentado en el libro de Russell titulado *Human Knowledge, Its Scope and Its Limits*, 1948) podía haber sido predicho, de hecho, yo lo había diagnósticado correctamente con mi observación en la Sociedad Aristotélica. Era una teoría de la inducción que aceptaba un principio inductivo —o unas reglas de inferencia inductiva— como válidas *a priori*. La diferencia entre el apriorismo de Russell y el de Kant estriba principalmente en la formulación de Russell de su principio inductivo como un conjunto de reglas de inferencia *probable*.

El método (trascendental), descrito por Russell en el pasaje que acabo de citar, equivale claramente a una renuncia a su enfoque subjetivista. Porque aquí acepta «el mundo de la física» como el hecho objetivo que la epistemología tiene que explicar. Así, incluso para Russell, el método subjetivo no es tan esencial como podría suponerse. Y no hay razón para que tenga que dominar los primeros pasos si se abandona posteriormente. El análisis del propio Russell muestra que la base subjetiva no puede fundamentar el realismo metafísico que él mismo quiere establecer y que se necesitan otros métodos —no subjetivos— para este fin.

---

[63] *Op. cit.*, p. 16. (El subrayado no está en el original.)

Esos otros métodos, sin embargo, no necesitan implicarnos a nosotros ni en el apriorismo de Russell ni en el de Kant. Aunque implican una ruptura con la filosofía empírica tradicional y especialmente con la metafísica de Berkeley y de Hume, que cuestiona la realidad del mundo físico, no nos obligan a romper con el empirismo, con la doctrina de que no puede establecerse como válido *a priori* ningún principio sintético.

Podemos combinar los dos, el empirismo y el realismo metafísico, siempre que nos tomemos en serio el carácter hipotético de todo «conocimiento científico» y el carácter crítico de toda discusión racional.

8. LA METAFISICA DE HUME. EL MONISMO «NEUTRAL»

Hume, como Russell, era un realista convencido a quien su teoría subjetiva del conocimiento llevó a resultados metafísicos que, aunque se sintió obligado a aceptar por motivos de lógica, era constitucionalmente incapaz de creer, ni siquiera durante una hora[64]. Hume parece haber despreciado su propia firme creencia en las cosas reales como una creencia irracional aunque prácticamente inevitable. Intentó hacer uso de esta misma contradicción —que él observaba en su propia mente— para resolver su problema; pero esto no le llevó a ningún lado: «La perplejidad que surge de esta contradicción», escribe, «crea una propensión a unir esas quebrantadas apariencias» —es decir, sus interrumpidas percepciones de un cuerpo— «por medio de la ficción de una existencia continuada...»[65]. Nadie que lea este tortuoso argumento (Libro I, Parte IV, sección II del *Treatise*) puede evitar el compartir su desilusión con los resultados finales de lo que con tanta confianza enuncia al principio como «mi sistema»[66]. Después de comprobar que incluso tomando al

[64] «...Sea cual fuere la opinión del lector en el momento presente, ...dentro de una hora estará convencido [de que] existe un mundo externo y uno interno». *Cf.* Hume, *Treatise*, fin de la Sección II del Libro I, Parte IV (Selby-Bigge, p. 218).

[65] *Treatise*, Selby-Bigge, p. 205. Véase también la nota en la p. 204 s.: «Este razonamiento, hay que confesarlo, es algo abstruso...; pero... esta dificultad misma puede convertirse en una prueba en favor del razonamiento...».

[66] *Treatise*, Selby-Bigge, p. 199, líneas 7-12. Parece que Hume cuando escribió este pasaje, al principio, no tenía intención de añadir a su sistema la «segunda parte» (p. 201); una «tercera parte de la hipótesis que me proponía explicar» (p. 205) y finalmente, incluso «un cuarto miembro de este sistema» (p. 208).

toro por los cuernos no podía hacerle moverse un paso más —que ni siquiera las contradicciones estimulaban a «la mente» a transcenderlas— declara francamente al final de la sección: «Sólo el descuido y la inatención pueden proporcionarnos algún remedio». Es extraño que este pensador tan claro y tan imparcial no sospechase nunca que al mostrar que su teoría no daba lo que prometía, le había dado lo que equivalía a un golpe mortal a su propia teoría subjetivista del conocimiento.

La reacción de Hume ante sus descubrimientos después de haber destruido efectivamente su propia teoría nos recuerda a la reacción de Frege ante la carta de Russell en la que éste le informaba de las paradojas que había descubierto en la teoría de Frege y suya. «Escribí a Frege», dice Russell, «que respondió con la mayor gravedad que *'die Arithmetik ist in's Schwanken geraten'*». Así pues, no era su teoría de la aritmética, sino la propia aritmética, pensaba Frege, la que estaba tambaleante y a punto de derrumbarse[67].

El «sistema» de Hume le obligaba al idealismo: la creencia en los cuerpos físicos —o en «objetos» que «continúan existiendo»— era para él la consecuencia de una inescapable «ficción de la imaginación»[68]. Estaba, pues, de acuerdo con Berkeley y «con los filósofos más sensatos», como lo expresa él, en que «nuestras ideas de cuerpos no son sino colecciones [de ideas] formadas por la mente...»[69]. Pero no sólo eran los cuerpos físicos los que quedaban así reducidos a simples manojos de ideas, sino también las mentes, incluida la de Hume. Porque Hume transcendió osadamente las concepciones de Berkeley enseñando que «...lo que llamamos *mente* no es más que un montón o colección de percepciones diferentes, ...a las que se suponen, aunque erróneamente, dotadas de una simplicidad e identidad perfectas»[70].

La ontología de Hume ha sido muy imitada, pero ninguno de los imitadores ha logrado del todo la hermosa simplicidad de su sistema. Esta puede ilustrarse convenientemente con el siguiente diagrama de una porción del universo —digamos que de un simposio en el que están presentes cinco filósofos (véase página 130).

Los círculos pequeños pueden interpretarse como visiones fugaces de cuerpos o como impresiones elementales o percepciones o como «datos sensoriales» o como su huella o reflejo de

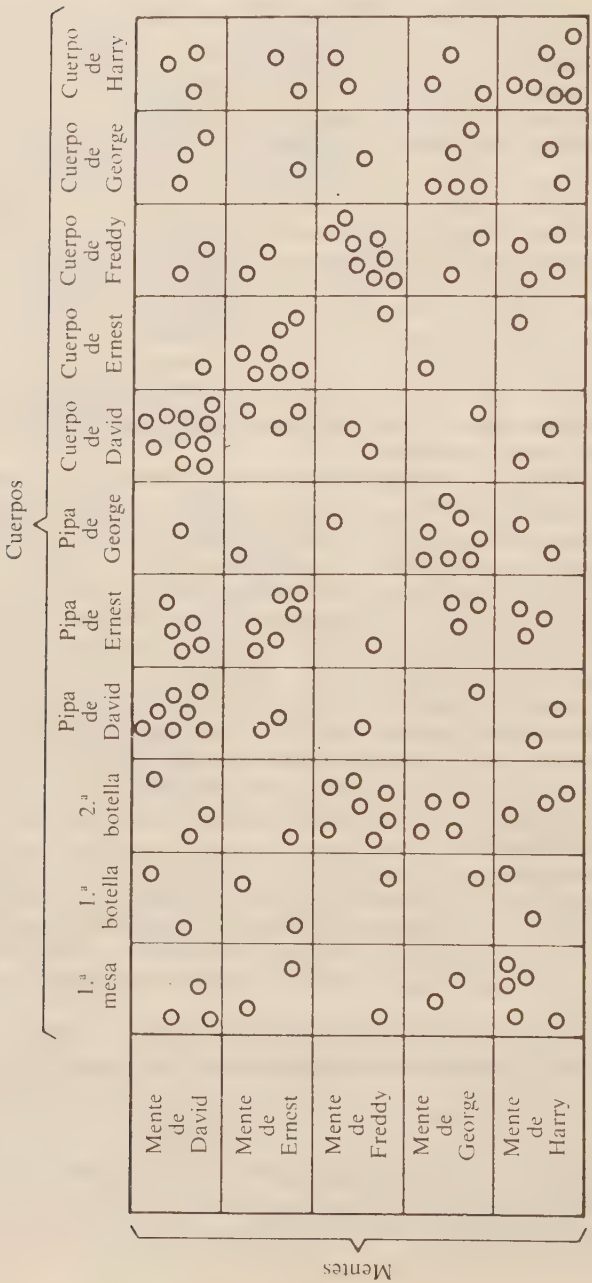
---

[67] Véase P. A. Schilpp, ed., *The Philosophy of Bertrand Russell, Library of Living Philosophers*, p. 13.

[68] *Op. cit.*, especialmente pp. 198 y 201.

[69] *Op. cit.*, p. 219.

[70] *Op. cit.*, p. 207; *cf.* también sección VI (p. 252, final).


Cuerpos
1.ª mesa
1.ª botella
2.ª botella
Pipa de David
Pipa de Ernest
Pipa de George
Cuerpo de David
Cuerpo de Ernest
Cuerpo de Freddy
Cuerpo de George
Cuerpo de Harry
Mentes
Mente de David
Mente de Ernest
Mente de Freddy
Mente de George
Mente de Harry

nuestra memoria. Ellos constituyen la materia prima de nuestro universo, según la teoría de Hume[71].

Cada uno de los manojos de percepción que constituyen cada uno de los *cuerpos* físicos a la vista, se representa aquí por medio de una columna vertical que contiene sus diversos aspectos tal como se les aparecen a las diversas mentes. Cada uno de los manojos de percepción que constituye cada una de las *mentes* de los asistentes al simposio está representado por una línea horizontal que contiene los diversos aspectos de los diversos cuerpos tal como se le aparecen a cada mente.

Puesto que el material, o los elementos, de universo de Hume consisten en percepciones de diversos cuerpos, pertenecientes a diversas mentes, este «sistema» es indudablemente idealista. Sin embargo, sistemas que son casi idénticos al de Hume —tales como los de Mach, o Russell o Reinchenbach— han sido descritos como sistemas de «monismo neutral», en donde se trata de dar a la palabra «neutral» el significado de «neutral entre el realismo y el idealismo» (o quizá incluso entre materialismo y espiritualismo) y donde «monismo» trata de indicar que el material del universo (o el carácter de sus elementos) es todo él de un tipo: sensaciones o impresiones[72].

Por lo que yo puedo comprender, la única diferencia entre el idealismo de Hume y el monismo «neutral» es ésta: si se le hubiera preguntado cómo están atados los diversos manojos, Hume hubiera respondido: por la *asociación de ideas*. Los «monistas neutrales» (Ernst Mach, William James o Bertrand Russell) habrían dado una respuesta diferente. Hubieran dicho (si yo les interpreto correctamente): «los manojos están ligados por *dos tipos de ley natural:* los manojos horizontales por leyes psicológicas (incluida la ley de asociación), y los manojos verticales por leyes físicas. En otras palabras, las mentes son aquellos manojos que obedecen las leyes de la psicología, mientras que los cuerpos físicos son aquellos manojos que observan las leyes de la física. Los elementos son al mismo tiempo elementos de cuerpos y elementos de mentes; son, pues, neutrales»[73].

Como el lector habrá supuesto, no puedo aceptar esta doctrina. Los elementos supuestamente «neutrales» son simple-

[71] [Compárese esta discusión con la de *El yo y su cerebro (S. & I. B.*, 1977), sección 53. *Ed.*]

[72] [Véase «Nota sobre Berkeley como precursor de Mach y Einstein» en *El desarrollo del conocimiento científico (C. & R.)*, capítulo 6, *Ed.*]

[73] Quizá pueda confesar aquí que en el invierno de 1926-7, yo mismo caí en esta forma de monismo neutral, sin darme cuenta, en aquel momento, de que era fundamentalmente lo mismo que la teoría de Mach y de James. Aunque abrigué la idea tentativamente durante un corto tiempo, para ver qué podía hacerse con ella, no creí seriamente en ella durante más de una hora, es decir, hasta que descubrí su carácter idealista.

mente percepciones, algo que puede hacerse desparacer cerrando los ojos o los oídos. Los cuerpos físicos, tal como los caracterizan las leyes de la física, no desaparecen de esta manera, mientras que las sensaciones visuales subjetivas sí que lo hacen. Así, el monismo «neutral» es una forma de idealismo; y sólo a un idealista —o a un filósofo totalmente imbuido de la teoría subjetivista del conocimiento— puede parecerle «neutral» esta doctrina.

Así pues, yo rechazo el «monismo neutral» junto con el idealismo, como partes de la teoría subjetivista del conocimiento.

9. POR QUE FRACASA LA TEORIA SUBJETIVISTA DEL CONOCIMIENTO

La teoría subjetivista del conocimiento falla por varias razones. Primero, supone ingenuamente que *todo* conocimiento es subjetivo, que no podemos hablar de conocimiento sin un conocedor, un sujeto que conozca. Segundo, lo que constituye tradicionalmente su problema fundamental está mal concebido. Estoy pensando en el problema (según la formulación de Russell; véase la nota 61 a las sección 7) «¿cómo sé yo lo que sé?» con su respuesta implícita, ingenuamente empirista: «a partir de la observación o de la experiencia de los sentidos»[74].

Contra todo esto, yo sostengo que el conocimiento científico no es ciertamente *mi* conocimiento. Porque resulta que yo sé lo poco que sé y sé cuántos miles de cosas hay que son «conocidas para la ciencia», pero no para mí (aunque me encantaría saberlas). Para mí (y espero que para cualquier otro sujeto) este solo hecho debería bastar para rechazar la teoría subjetivista del conocimiento científico.

Pero incluso esas migajas de conocimiento científico y de sentido común que poseo yo mismo, no se conforman al esquema preconcebido de la teoría subjetivista del conocimiento: pocas de ellas son enteramente consecuencia de *mi propia* experiencia. Más bien, en gran medida, son el resultado de que yo haya absorbido ciertas tradiciones (por ejemplo, leyendo ciertos libros), en parte consciente y en parte inconscientemente. Y no están más estrechamente vinculadas a mi propia

[74] [En la obra posterior se desarrolla y extiende considerablemente el argumento contra la teoría subjetivista del conocimiento. Véase *Conocimiento Objetivo (Obj. Kn.)*, capítulos 1-4; *Búsqueda sin término (U. Q.)*, secciones 31, 38 y 39; y *El yo y su cerebro (S. & I. B.*, con Sir John Eccles), capítulo P3. *Ed.*]

experiencia observacional de lo que lo están mis creencias metafísicas (convicciones religiosas o morales, por ejemplo) que son también consecuencia de la absorción de ciertas tradiciones. En ambos casos, mi propia crítica de algunas de esas tradiciones puede jugar un papel importante en la formación de lo que yo creo que sé. Pero esta crítica está casi siempre estimulada por el descubrimiento de inconsistencias dentro de una tradición o entre tradiciones diferentes. (Casi nunca me estimula a la crítica el descubrimiento de una inconsistencia entre una tradición y una experiencia observacional mía; porque a muy pocos se les concede el falsar realmente, por medio de sus propias observaciones, una teoría tradicional.)

Así pues, el conocimiento científico no es idéntico a *mi* conocimiento; y lo que constituye *mi* conocimiento —*mi* conocimiento de sentido común o *mi* conocimiento científico— es en gran medida el resultado de mi asimilación de tradiciones y (espero) de algo de pensamiento crítico.

Hay, naturalmente, un tercer tipo de conocimiento que también podría llamarse «mío»: sé dónde tengo que buscar mi tintero o la puerta de mi habitación; conozco el camino a la estación de ferrocarril; sé que los cordones de mis zapatos tienen una tendencia a romperse si voy retrasado. Este tipo de conocimiento (que podría denominarse «conocimiento personal»), no es en absoluto tradicional, puesto que es consecuencia de *mis propias experiencias* y, por tanto, es el que más se acerca al tipo de conocimiento contemplado por la teoría subjetivista. Pero ni siquiera este «conocimiento personal» encaja en dicha teoría; porque está engastado en el conocimiento de sentido común de las cosas tradicionales: de tinteros, cordones de zapatos, estaciones de ferrocarril; cosas sobre las que tenemos que aprender absorbiendo una tradición. Admitámoslo, nuestras observaciones, nuestros ojos y oídos, fueron una ayuda inmensa en este proceso de absorción. Pero absorber una tradición es un proceso fundamentalmente diferente del contemplado por la teoría subjetivista que quiere que yo comience a partir de *mi* conocimiento y, además, a partir de mi experiencia observacional.

Es probable que el subjetivista replique que los procesos que él contempla son lógica y genéticamente anteriores a aquellos a los que yo estoy aludiendo. Dirá que antes de que yo pueda empezar a absorber una tradición, tengo que aprender mucho sobre el mundo por medio de la experiencia observacional.

En este punto quiero preguntar, como un subjetivista: «¿Cómo lo sabe usted? ¿Cuál es la base observacional de su aserción?» Porque su réplica es consecuencia de meros prejuicios y falta de imaginación. Las ciencias biológicas nos cuentan

una historia diferente: por ejemplo, tenemos muy buenas razones para pensar que muchos insectos tienen mucho «conocimiento innato», que algunos, de hecho, son incapaces de aprender, de modificar sus reacciones innatas, en ciertos terrenos; que usan sus sentidos y «reconocen» ciertas cosas (el alimento o una posible pareja), sin haberlas visto antes. Pero si esto es verdad respecto a algunos insectos, puede muy bien ser verdad respecto a los mamíferos, al menos hasta cierto punto, y respecto a los hombres: podemos muy bien nacer con algo parecido a tradiciones innatas o «instintos» (aunque éstos nos engañen a menudo). De hecho, la respuesta del subjetivista es una parte de la mitología empirista.

La siguiente respuesta del subjetivista será probablemente ésta: incluso si fuera verdad que los insectos y los hombres tienen conocimiento innato de un tipo instintivo, este conocimiento seguiría teniendo que ser el resultado, de una u otra forma, de experiencia observacional; es decir, de la experiencia de generaciones anteriores.

Por dos razones diferentes esta respuesta es tan insostenible como la anterior. En primer lugar, incluso si fuese verdad, cambiaría la cuestión. En segundo lugar, lo que afirma parece ser falso y, en cualquier caso, choca con las concepciones que actualmente aceptan la mayor parte de los genéticos.

En cuanto al primero de mis puntos, lo que está en discusión no es el empirismo en general, sino el empirismo subjetivista o, más exactamente, la teoría subjetivista del conocimiento que parte del problema de «¿cómo sé yo?» y que supone que el origen de todo conocimiento puede encontrarse en mis experiencias observacionales subjetivas. Pero una vez que se admite que *mis* experiencias no bastan, que tenemos que acudir a las de mis antepasados (o a las de cualquier otro), la teoría subjetivista se desmorona. Porque se admitiría entonces que el «conocimiento» es algo inter-subjetivo y entonces se admitiría que este conocimiento intersubjetivo tendría que ser *genéticamente* anterior a mi conocimiento subjetivo. (Esto, por cierto, demuestra *a fortiori* que el último no puede ser *lógicamente* anterior al primero.)

Respecto al segundo punto, la réplica es esencialmente lamarckiana, en oposición a una respuesta darwiniana; porque esencialmente se intenta que sea una defensa de la concepción según la cual si tenemos conocimiento innato, debe, en último término, seguir siendo el resultado de *experiencia observacional adquirida individualmente*. Yo, personalmente, tengo simpatía por el lamarckismo, pero el lamarckismo no es lo que apoyan actualmente la mayoría de los biólogos; y, en cualquier caso, una teoría del conocimiento (sea empirista o de otro tipo) no

debería adoptar acríticamente una conjetura tal como el lamarckismo basándose en algo parecido a argumentos *a priori*[75].

A los subjetivistas suelen inducirles a error con frecuencia algunas de las ambigüedades de la palabra «conocimiento». «Conocimiento» se deriva claramente de «yo conozco». Esto sugiere que el «conocimiento» sólo puede ser aquello que es conocido: conocido para las personas.

Pero esta concepción es claramente inadecuada. Tomemos un libro que contenga una tabla de logaritmos. Hay personas que saben cómo hacer esa tabla (cómo calcularla, cómo ordenarla, cómo imprimirla), y otras que saben cómo usarla; pero no hay nadie que «sepa» la tabla (ni siquiera en el sentido, por ejemplo, en el que algunas personas «saben» el principio de la tabla de multiplicar). Sin embargo, la tabla representa «conocimiento», conocimiento objetivo: respuestas, o respuestas parciales, a incontables cuestiones importantes: una información valiosísima. Y este conocimiento no es «conocido» por nadie (ni siquiera por el compilador): está sólo *disponible;* está ahí, potencialmente, para todo el que esté dispuesto a confiar en el redactor y el editor.

El caso es similar con todas las teorías científicas: una teoría puede contener, potencial o disposicionalmente, una riqueza de información que nadie «conoce», ni su descubridor ni sus usuarios; puede ser utilizada y se puede obtener información de ella; por ejemplo, aplicándola a conjuntos de condiciones muy especiales.

El conocimiento científico, en este sentido objetivo, puede ser estudiado, asimilado, usado, aplicado. Una y la misma parte de él puede ser aceptada o rechazada, dogmática o críticamente. Puede ser creída fervientemente o considerada como una conjetura aproximada: hay muchas actitudes subjetivas divergentes y muchas formas de reaccionar ante las tradiciones.

Uno de los usos del conocimiento objetivo, o de una de las partes del conocimiento objetivo, es ayudarnos a formar algunas de nuestras propias convicciones subjetivas. Este camino —el camino que lleva desde el conocimiento objetivo al conocimiento subjetivo— se toma con mucha mayor frecuencia que el camino que lleva del conocimiento subjetivo al conocimiento objetivo. Sin embargo, la teoría subjetivista del conocimiento supone, como cuestión de hecho, o bien que *sólo* existe conocimiento subjetivo (que un libro es un mero objeto físico de un tipo que puede estimular en la mente de su usuario aquellas asociaciones que pueden constituir conocimiento) o al menos

[75] [Véase *Conocimiento Objetivo (Obj. Kn.)* pp. 87, 139, 244-59. *Ed.*]

que el conocimiento objetivo —si podemos hablar en absoluto de tal cosa— no puede ser otra cosa que el registro o una derivación de alguna parte de conocimiento subjetivo.

Aunque el conocimiento objetivo siempre es el resultado directo o indirecto de acciones humanas, de pasos dados a la luz del conocimiento subjetivo y objetivo, el conocimiento objetivo emerge a menudo sin que previamente haya sido conocido subjetivamente. Este es el caso, invariablemente, en todos los cálculos (en la medida en que atañen al hombre que los hace): aquí esperamos a que surja el resultado en alguna forma física antes de formarnos la correspondiente convicción subjetiva. (No deberíamos olvidar que los inventores del cálculo no sabían realmente cómo funcionaba, aunque tenían una vaga idea, de modo que obtuvieron su conocimiento a partir de sus cáculos casi por arte de magia. Lo que descubrieron fue que la máquina del conocimiento que habían construido funcionaba extremadamente bien, por lo general, aunque no siempre.)

Se verá, por lo que he dicho, que podemos considerar el conocimiento objetivo —la ciencia— como una *institución social,* o un conjunto o estructura de instituciones sociales[76]. Como otras instituciones sociales, es el resultado, en gran medida no intencionado y casi enteramente imprevisto (*pace* Bacon) de acciones humanas. Desde luego, vive y crece en gran medida a través de la cooperación y competición institucionalizadas de los científicos, que no están inspirados solamente por la curiosidad —el deseo de aumentar su conocimiento subjetivo—, sino en mayor grado por el deseo de *hacer una contribución al aumento del conocimiento,* esto es, del conocimiento objetivo. (Muchas de las grandes contribuciones que se le han hecho consisten en errores y en la detección de esos errores.)

El estudio de cómo se hacen las contribuciones, de cómo se contrastan, aceptan, rechazan; de su condición de hipótesis, de las normas tradicionales que se les aplican y del refinamiento de esas normas, este estudio, es la parte más interesante y más fructífera de la teoría del conocimiento —de la teoría del conocimiento objetivo, naturalmente—.

No pretendo, desde luego, sugerir que no existe el conocimiento subjetivo; o que no es importante para el aumento del conocimiento objetivo: el conocimiento objetivo no podría aumentar si él. La relación entre los dos no es, como he indicado, nada sencilla. Por tanto, hay razones muy importantes para tener también una teoría del conocimiento subjetivo.

---

[76] Véase *La miseria del historicismo (The Poverty),* especialmente secciones 32 y 21 y *La sociedad abierta (O.S.),* esp. capítulos 23 y 14.

Sin embargo, esa teoría formaría parte de una ciencia empírica, y no de la lógica de la ciencia o epistemología. Porque su objeto es el aumento, el desarrollo, del conocimiento de alguien, el aumento de esas experiencias subjetivas que expresamos cuando decimos: «¡Sabía que era así!». Pueden convertirse en el objeto de un estudio psicológico o —de mayor interés, pienso— de un estudio biológico general.

Si queremos considerar el conocimiento y la actividad científica como fenómenos biológicos, tendremos que tener en cuenta su papel en el proceso de adaptación por el animal humano a su entorno y a cambios ambientales: a los sucesos que se ciernen sobre ellos. Los animales e incluso las plantas se preparan para la hibernación, pongamos por caso, de una manera puramente instintiva (¿cómo sé esto y las demás cosas que voy a afirmar aquí? No las sé: todo esto son conjeturas) y algunas personas lo hacen, quizá, consultando el calendario y haciendo un encargo al carbonero. Unas cuantas más lo hacen estudiando la combustión del petróleo e inventando un nuevo tipo, más seguro y más económico, de estufa de petróleo.

El análogo psicológico o biológico de una hipótesis puede describirse como una *expectativa o la espera* de un suceso. Esta expectativa o espera puede ser consciente o inconsciente. Consiste en la aptitud de un organismo para actuar o reaccionar en respuesta a una situación de un cierto tipo específico. Consiste en la activación (parcial) de ciertas *disposiciones* (véase la sección 27 de *L. I. C.*, especialmente la p. 94, donde se trata brevemente del carácter disposicional del conocimiento).

Los ejemplos clásicos del modo en el que las expectativas inconscientes pueden convertirse en conscientes son: dar un traspié al bajar un escalón («pensé que no había escalón aquí») u oír pararse el reloj («no me daba cuenta de estar oyendo el tic-tac, pero lo he oído pararse»). Nuestro organismo esperaba, inconscientemente, ciertos sucesos y sólo nos hicimos conscientes del hecho cuando nuestras expectaciones resultaron decepcionadas o falsadas.

Esta disposición a estar preparado para lo que va a venir parece ser el verdadero análogo biológico del conocimiento científico.

En un organismo animal, las disposiciones a reaccionar de cierta manera ante ciertos tipos de estímulos son, en parte, innatas. Mi tesis es que, en la medida en que son *adquiridas*, son *modificaciones* de disposiciones innatas que son «plásticas» y que se desarrollan y cambian, al ser activadas por estímulos y sobre todo bajo la influencia del fracaso y el éxito (unidos probablemente a sentimientos de dolor o de placer); porque las acciones y reacciones liberadas por los estímulos están dirigi-

das, por lo general, hacia ciertas metas biológicas. De este modo, el organismo desarrolla su conocimiento disposicional innato: aprende por el método de prueba y error.

Este es conocimiento adquirido que puede transmitirse por tradición, aunque no (si debemos creer la conjetura darwiniana) por herencia. Sin embargo, también el conocimiento innato cambia; y la supervivencia del conocimiento innato más apto es, una vez más, un proceso de prueba y error[77].

El cuadro aquí representado debería compararse con el de Hume, que sigue teniendo éxito entre los epistemólogos: la concepción de que, de hecho, aprendemos *por medio de la repetición de observaciones*.

Es cierto que la repetición juega un papel en el proceso de aprendizaje. Pero éste es aprendizaje en un sentido totalmente diferente del aprendizaje por prueba y error (véase la sección 3). Si aprendemos un arte (montar en bicicleta, tocar el piano, hablar un idioma nuevo), entonces los dos procesos pueden parecer uno, porque el aprendizaje por repetición empieza donde termina el aprendizaje por prueba y error. Pero ésta es la diferencia: todo *nuevo* conocimiento, todas las modificaciones radicales de nuestras disposiciones, todos los descubrimientos, son el resultado de pruebas y errores. La repetición no hace más que familiarizarnos con el conocimiento recién adquirido: puede hacernos olvidar lo que estamos haciendo y cómo hemos adquirido nuestras reacciones y, sobre todo, que había que vender dificultades durante el período de prueba y error.

Que yo sepa, los teóricos del proceso de aprendizaje no tienen muy claro el abismo que separa a estos dos sentidos de «aprendizaje». La teoría del conocimiento subjetivo —del aumento de nuestro conocimiento subjetivo— cometió su equivocación fatal cuando consideró la repetición en vez del proceso de prueba y error como el principal instrumento del aprendizaje. (Véase la sección 3, más arriba. Esta equivocación juega también un papel decisivo en todas las teorías subjetivistas de la probabilidad: véase la parte 2 de este tomo.)

Respecto a la observación y el aprendizaje a través de la evidencia de nuestros sentidos, la teoría aquí esbozada está en directa oposición al empirismo tradicional.

La teoría tradicional está excesivamente influida por la innegable verdad de que sin nuestros sentidos no podríamos tener ningún conocimiento del mundo. Estoy dispuesto a admitir más aún: sin nuestros sentidos, no podríamos vivir. Pero

[77] [Véase Donald T. Campbell, «Evolutionary Epistemology», en Paul A. Schilpp, ed., *The Philosophy of Karl Popper* (La Salle, 1974), pp. 413-463. *Ed.*]

estoy en total desacuerdo con la concepción empirista tradicional de que todo nuestro conocimiento del mundo es el «resultado» de la observación o de la experiencia de los sentidos o que el conocimiento siempre «entra en nuestro intelecto a través de nuestros sentidos». Aparte del conocimiento disposicional innato, nuestro conocimiento se desarrolla, efectivamente, bajo la influencia de *estímulos* (transmitidos por nuestros sentidos). Pero los estímulos, por regla general, actúan como meros *gatillos*, aunque puedan desatar torrentes de nuevas tendencias en nuestro equipamiento disposicional. Lo mismo ocurre con las observaciones. Son significativas sólo en el contexto de nuestras expectativas, nuestras hipótesis, nuestras teorías.

La concepción tradicional de que nuestro conocimiento aumenta por la acumulación (o quizá por la repetición) de percepciones y observaciones es puro mito, probablemente el mito más extensamente creído en los tiempos modernos. Lo refuta el hecho de que un hombre ciego y sordo puede saber más y hacer mayores contribuciones al conocimiento que uno con una vista y un oído muy agudos: ciertamente, si nuestros sentidos fuesen intelectualmente tan cruciales como creían los empiristas, entonces el fallo de esos sentidos tan importantes produciría una deficiencia intelectual de la mayor gravedad. Pero nada de eso ocurre, como lo muestra el gran ejemplo de Helen Keller.

El empirismo tradicional trata de describir la mente, con la ayuda de metáforas, con una *tabula rasa* —algo parecido a una pizarra bien limpia o a una placa fotográfica no expuesta— que debe ser grabada mediante las observaciones. Esta teoría, a la que he llamado *«la teoría de la mente como cubo»*, ve la mente como un cubo y los sentidos como embudos a través de los cuales el cubo puede ir llenándose lentamente de observaciones. La suma total de esas observaciones (o quizá la suma total ordenada o digerida) es «nuestro conocimiento». Esta concepción es radicalmente errónea.

Hay dos aspectos, por así decirlo, en cada organismo superior: su constitución innata, sus disposiciones a actuar y a reaccionar, su «reactividad», es uno de ellos, y su aparato para recibir estímulos, su «sensibilidad» es el otro. Mientras el empirismo tradicional considera que el conocimiento está localizado en la esfera de la *sensibilidad*, mi concepción lo sitúa dentro de la esfera de la *atividad y reactividad* del organismo.

Este cambio tiene consecuencias de largo alcance. Según mi concepción, las observaciones (o «sensaciones» o «datos sensoriales», etc.) no son nada parecido a las «uvas» de Bacon, de las cuales fluye el «vino del conocimiento»: no son la materia prima del conocimiento. Por el contrario, las observaciones

presuponen siempre un conocimiento disposicional previo. Una observación es el resultado de un estímulo que *toca una campana*. ¿Qué significa esto? El estímulo debe ser *significativo*, relativo a nuestro sistema de expectativas o esperanzas, para ser capaz de tocar una campana y, por tanto, ser observado. La campana —o, en otras palabras, un interés, una disposición a responder al estímulo, una expectativa de que el estímulo resulte ser significativo— tiene que estar allí de antemano, o si no el estímulo pasará inadvertido: no estimulará. Así pues, las observaciones presuponen una preparación disposicional, es decir, un conocimiento disposicional; y aunque las observaciones puedan desatar cambios disposicionales —especialmente si son contrarias a nuestras expectativas— no pueden de ningún modo convertirse en partes o ingredientes del sistema de disposiciones que constituye nuestro conocimiento subjetivo: no son nunca afines al conocimiento subjetivo (aunque sus *resultados* pueden serlo), sino que pertenecen a un dominio diferente de cosas.

Por tanto, la manera habitual de considerar un «reflejo condicionado» es totalmente errónea. Lo que hacen los experimentos no es «vincular el sonido de una campana con el flujo de saliva del perro»; más bien se aprovechan del hecho de las reacciones, vitalmente importantes, conectadas con los esfuerzos del perro para conseguir su alimento, constituyen, por razones obvias, uno de los sistemas disposicionales más plásticos del animal: los animales que cazan para obtener comida en condiciones muy diferentes tienen que adaptarse a esas condiciones. La obtención del alimento es, pues, un terreno predispuesto para el aprendizaje: la plasticidad misma es innata. De ahí que el perro pueda formarse fácilmente una nueva expectativa, de hecho, una nueva teoría: descubre que la campana le anuncia la cena. Eso es todo. Cuando, desde un principio, las expectativas son menos plásticas, o cuando los intereses vitales del animal no están implicados, los intentos de establecer un reflejo condicionado fallan generalmente.

El empirismo tradicional supone que podemos obtener o extraer nuestro conocimiento del mundo a partir de nuestra experiencia sensorial por métodos que pueden considerarse razonables. Si se acepta acríticamente esta concepción —la teoría del conocimiento como cubo—, entonces el descubrimiento de Hume de que es imposible justificar la inducción debe conducir al irracionalismo. Debe llevar a la conclusión de que, por ejemplo, nuestra creencia en los cuerpos físicos es un prejuicio injustificable. Y, en efecto, lo es. Como todo nuestro conocimiento, es, en principio, un prejuicio, pero uno que puede ser examinado críticamente. Como resultado del examen crítico de

esta conjetura, que se extiende a lo largo de los últimos cien años, sabemos ahora más sobre los cuerpos, por ejemplo, sabemos que son procesos (como previó Heráclito). Sin embargo, procedemos de un modo perfectamente racional: aprendemos, extendemos nuestro conocimiento, *contrastando* nuestros prejuicios, por el método de prueba y error antes que por inducción mediante repetición. Ni siquiera nuestra «fe animal» en la regularidad es el resultado de las repeticiones. Proviene de una disposición innata (que puede ser activada por el estímulo de las repeticiones o por el de un suceso único) a esperar regularidades.

Así pues, el problema de la teoría subjetivista del conocimiento, «¿cómo aprendo, a través de mis sentidos, a saber sobre el mundo?» está mal planteado. Y la pregunta de Hume «¿Qué causas nos inducen a creer en la existencia del cuerpo?»[78] puede responderse de una forma radicalmente diferente de como lo hizo Hume. Podemos decir que un perro «cree que es existente» aquello a lo que gruñe o muerde o aquello que le gusta. Existentes, en el sentido que Hume tiene en mente, son, en el más alto grado, aquellas cosas ante las que reaccionamos con fuerza: contra las que luchamos por defender nuestras vidas, o las que comemos o por las que podemos resultar aplastados. La creencia en su existencia proviene de una disposición innata a tratarlas como importantes en potencia. Es una creencia innata, aunque sin duda necesita de muchos estímulos para desarrollarse. Estas son las causas —al menos según mi conjetura— de esta creencia, de este prejuicio. Pero no son razones. Las razones sólo empiezan a aparecer cuando empezamos a hacer la crítica de nuestros prejuicios, a tratarlos como conjeturas; por ejemplo, cuando empezamos a descubrir más cosas sobre los «cuerpos». Y puede ser que aparezcan más razones de otro tipo cuando critiquemos la epistemología subjetivista que llevó a la pregunta de Hume.

La suposición dogmática de que debemos ser capaces de responder a una pregunta sobre la teoría del conocimiento tal como la de Hume por medio de una investigación sobre el modo en que surge una creencia a partir de percepciones y asociaciones o sobre el modo en que «experimentamos inmediatamente» las manchas de color o las cosas, etc., está en la raíz de la mayoría de las formas de idealismo. Es sorprendente que, cien años después de Darwin, los filósofos continúen discutiendo ingenuamente el problema de la epistemología en términos del origen de nuestro conocimiento en datos sensoriales o percepciones (o en términos de los tipos de palabras que usa-

[78] *Treatise*, Libro I, Parte IV, Sección II, primer párrafo.

mos cuando discutimos sobre percepciones) o en términos del número de «repeticiones» de la «observación» de un cuervo negro o de un cisne blanco[79].

El motivo más profundo tras la teoría subjetivista del conocimiento es la comprensión de que nuestro «conocimiento» es incierto (y, por tanto, no es realmente «conocimiento») y el deseo de empezar a partir de la *certeza:* de una base *cierta* o al menos de la base más cierta que tengamos. Y las experiencias que «me son dadas» tales como la experiencia de ver algo (o, con Descartes, la experiencia de dudar) parecen ofrecerse como los puntos de partida naturales. Los subjetivistas suponen acríticamente que sobre la base de esos «datos» puede erigirse el edificio del conocimiento —del conocimiento científico—. Pero este supuesto es incorrecto. No puede construirse nada sobre esos «datos», incluso aunque supongamos que existen. Pero no existen: no existen los «*datos*» sin interpretar; no hay nada que nos sea dado puramente, sin interpretar; nada que tomar como base. *Todo* nuestro conocimiento es interpretación a la luz de nuestras expectativas, nuestras teorías, y es, por tanto, *hipotético* de una u otra manera.

Esto es exactamente lo que debemos esperar si el realismo es verdadero, si el mundo a nuestro alrededor es, más o menos, como el sentido común refinado por la ciencia nos dice que es. Si el realismo es verdadero, si somos animales tratando de adaptarnos a nuestro medio, entonces nuestro conocimiento no puede ser más que el asunto de prueba y error que he descrito. Si el realismo es verdadero, nuestra creencia en la realidad del mundo y en las leyes físicas *no puede* ser demostrable, no puede mostrarse que es seguro o «razonable» por medio de ningún razonamiento válido. En otras palabras, si el realismo está en lo cierto, no podemos esperar tener *más que* un conocimiento conjetural: el milagro es, más bien, que hayamos tenido tanto éxito en nuestra búsqueda de conjeturas. Y al decir esto, no tengo en mente sólo los milagrosos éxitos de los últimos trescientos años.

Por tanto, si el realismo está en lo cierto, el objetivo de la teoría subjetivista, de encontrar una base subjetiva segura sobre la que erigir nuestro conocimiento del mundo —y sólidas razones para una creencia en la realidad del mundo— es un objetivo irrealizable y, en realidad, nada razonable.

---

[79] Véase también mi artículo «Philosophy of Science: A Personal Report», especialmente las secciones IV y V, y la segunda nota a pie de página en la sección II. [Este artículo se ha reimpreso como Capítulo 1 de *El desarrollo del conocimiento científico (C. & R.)*. Véanse también los capítulos 1, 2, 3 y 7 de *Conocimiento Objetivo (Obj. Kn.*, 1972). *Ed.*]

El realismo, pues, nos explica por qué la situación de nuestro conocimiento es *necesariamente* precaria. Si, por otro lado, alguna forma de idealismo es verdadera, entonces puede ocurrir *cualquier cosa* y, por tanto, *posiblemente*, también lo que de hecho ocurre. Así, el realismo es la más fuerte de las dos teorías metafísicas, desde un punto de vista lógico. Es preferible por razones lógicas: el idealismo metafísico resulta estar falto de *todo* poder explicativo.

Al rechazar la teoría subjetiva del conocimiento debilitamos uno de los más fuertes argumentos o motivos del idealismo. Pero aún tenemos que discutir los argumentos escépticos de Hume contra la realidad de la materia.

10. UN MUNDO SIN ENIGMAS

La teoría subjetiva del conocimiento —que nos dice que construyamos el mundo físico a partir de mi propia «egocéntrica» experiencia perceptual— se ha fijado a sí misma una tarea que es al mismo tiempo innecesaria y de una dificultad que la hace imposible. Esta es la razón de que siempre caiga en algún tipo de idealismo. Y lo que hace tan poco atractivo al idealismo es, precisamente, la fatal facilidad con la que lo explica todo. Porque el idealismo resuelve todos los problemas: vaciándonos de contenido.

Un ejemplo típico es la solución del idealista (o del «monista neutral») al problema del cuerpo y de la mente —el problema de las inmensamente intrincadas influencias psicológicas (de las drogas, por ejemplo) sobre nuestro estado mental y, viceversa, el problema de las influencias mentales (por ejemplo, del darse cuenta de los peligros) sobre nuestro estado fisiológico—. El idealismo en todas su formas, incluido el monismo neutral, nos permite resolver este problema en la fracción de un segundo, tan completamente que desaparece. Si todos los cuerpos, incluyendo el mío, no son más que manojos de percepciones, entonces, los cambios en el mundo físico serán, en general, también cambios en el mundo mental y viceversa: puede predecirse así que *habrá interrelaciones* (cuyas leyes detalladas habrá que descubrir por inducción).

Otro problema que se esfuma a la luz del subjetivismo y del idealismo —o quizá en la oscuridad que ellos crean— es el problema de la materia. Este es uno de los problemas más antiguos de la filosofía. Está estrechamente relacionado con el problema del cambio. El cambio parece presuponer la existencia de algo que permanece inmutable durante el cambio. Esto llevó a Heráclito y a Hegel a la teoría de la identidad de los

contrarios; y llevó a Leucipo y a Demócrito a la teoría, más importante, de que *todo cambio consiste en movimiento:* de partículas materiales indestructibles, de átomos que no están sujetos a ningún cambio distinto del movimiento.

Esta era una teoría clara e importante del cambio, y de la materia: la materia era esa entidad que podía moverse pero que, de otro modo, permanecía inmutable durante el cambio. Ni esta teoría del cambio ni esta teoría de la materia fueron definitivas, naturalmente. Desde los tiempos de Leucipo y Demócrito y, más especialmente, desde Descartes —con Leibniz, Boscovich y Faraday— la teoría de la materia ha evolucionado hasta convertirse en una teoría de la *estructura de la materia misma*; una evolución que constituye uno de los capítulos más fascinantes de la historia del pensamiento científico y filosófico (véase el tomo III del *Post Scriptum*, sección 20)[80].

Pero, ¿qué tiene la epistemología subjetivista que decir de todo esto? «Cuando seguimos paso a paso un objeto a través de sus cambios sucesivos, el progreso continuo del pensamiento nos hace atribuir una identidad a la sucesión», escribe Hume, «cuando comparamos su situación después de un cambio considerable, el progreso del pensamiento está interrumpido; y, en consecuencia, se nos presenta la idea de diversidad: para reconciliar estas contradicciones, la imaginación tiende a idear algo desconocido e invisible de lo que supone que permanece igual durante todas esas variaciones; y a este algo ininteligible, lo llama *una sustancia, o... materia*»[81].

La teoría de la sustancia o materia está descrita de modo inteligible en este pasaje, algo difícil, de Hume, y su conexión con la teoría del cambio se indica claramente; pero se desecha como una ficción ininteligible de «*la imaginación*». (Como este *deus ex machina* pertenece a la mente, no debe ser «mas que» un manojo de ideas; y las ideas no son ininteligibles. Es raro que Hume nunca se sintiera molesto con *esta* contradición.) Además el pasaje indica con bastante consistencia que los objetos no observados son, igualmente, sólo ficciones de la imaginación; son interpolaciones ficticias entre los fenómenos observados: son «*interfenómenos*» por usar un término moderno que debemos a Reichenbach (véase la sección 13, más adelante).

Esta «ficción ininteligible» de la sustacia o la materia no es una idea, según Hume: tampoco es una percepción o una impresión ni un reflejo; y, por tanto, carece de sentido. Porque Hume silogiza: «No tenemos una idea perfecta de nada, excepto de una percepción. Una sustancia es enteramente dife-

[80] [Véase *El yo y su cerebro (S. & I. B.)*, capítulos P1 y P3. *Ed.*]
[81] *Treatise*, Libro I, Parte IV, Sección III (Selby-Bigge, p. 220).

rente de una percepción. Por tanto, no tenemos ninguna idea de una sustancia»[82]. Y aún más, concluye de esto que cualquier cuestión referente a una sustancia debe carecer de sentido: «¿Qué posibilidad hay, pues, de responder esta pregunta... cuando no... conocemos el sentido de esta pregunta?»[83].

Así pues, el problema de la materia, y con él, el de la estructura de la materia, se descarta como carente de sentido, con unos argumentos típicamente *a priori*. Hume siguió aquí a Berkeley que había descartado la teoría atómica como carente de sentido[84] y fue seguido, a su vez, por Mach.

Esta supresión de lo que estaba destinado a convertirse en uno de los problemas más importantes de la ciencia y de la filosofía no fue accidental. En el mundo de Berkeley, Hume y Mach no puede haber problemas serios. Según el positivismo «nuestro mundo es sólo superficie: no tiene profundidad»[85]. Porque no consiste más que en nuestras percepciones y en sus reflejos en nuestra memoria. Es un mundo en el que no hay nada que descubrir, puesto que nada está cubierto. Es un

---

[82] *Treatise*, Libro I, Parte IV, Sección V (Selby-Bigge, p. 234). Sin duda, la primera premisa de Hume debía continuar «y ninguna de nuestras ideas (perfecta o no) es enteramente diferente de una percepción». Pero nada, creo, puede salvar el silogismo. Cambiando el orden de las premisas de Hume, puede expresarse: «Una sustancia es algo enteramente diferente de una percepción. Ninguna de nuestras ideas es enteramente diferente de una percepción. Luego, ninguna de nuestras ideas es la idea de una sustancia». Pero la conclusión correcta diría: «Luego ninguna de nuestras ideas es una sustancia», lo que sería trivial, considerando que Hume ha descrito la sustancia como comprendiendo esencialmente ideas *e* interfenómenos. Para obtener la conclusión de Hume, su segunda premisa, que es la primera de las nuestras, tendría que sustituirse por «la idea de la sustancia es algo enteramente diferente de una percepción»; pero esto sería falso *a priori*, según Hume, puesto que según su concepción, *ninguna* idea es enteramente diferente de una percepción. La dificultad fundamental, puesto que en un mundo que no contiene nada más que ideas (y en el cual una creencia es meramente «una idea vivaz») no podemos explicar nuestra creencia en la sustancia como nada parecido a una «ficción» que *no sea una idea*. Berkeley y los modernos analistas del lenguaje hubieran evitado esta dificultad diciendo que la *palabra* «sustancia» carece de significado (véase también *Treatise*, Selby-Bigge, pp. 61-62); pero entonces surge la dificultad bajo una nueva forma: hasta ahora no se ha formulado todavía ningún criterio sobre el significado de las palabras que tenga sentido. Pero, a diferencia de los analistas del lenguaje, Hume no estaba realmente interesado en las palabras sino en los mundos: los mundos de nuestras ideas y de nuestras creencias.

[83] *Treatise, loc. cit.* Con todo, Hume parecía entender el significado de la pregunta de si una sustancia podía, o no, ser una de nuestras ideas, y pensaba que podía responderla.

[84] Véase mi «Nota sobre Berkeley como precursor de Mach», ahora en *El desarrollo del conocimiento científico (C. & R.)*, capítulo 6.

[85] Aquí cito de memoria y, probablemente, no exactamente, un brillante comentario en una conversación con Otto Neurath, cuando resumía su propia filosofía de la naturaleza.

mundo sobre el que no hay nada que averiguar, nada que aprender. Es un mundo sin enigmas.

Así, la epistemología de Hume, aparentemente tan escéptica, lleva, como la de Bacon, a la doctrina de que la naturaleza es un libro abierto y que la verdad está manifiesta: como no hay nada más que ideas, no puede haber enigmas. Esto explica por qué la prueba de Hume de la invalidez de la inferencia inductiva se tomaba tan raras veces en serio (al contrario que su crítica de la causalidad). Si la naturaleza fuese un libro abierto, no necesitaríamos una lógica inductiva: podríamos inducir, de forma no válida, pero con bastante éxito, por asociación o por costumbre.

Hume podría muy bien haber considerado que comentarios como éste constituían una seria crítica de su obra: «Pues, si la verdad está, en absoluto, al alcance de la capacidad humana, es seguro que debe hallarse muy profunda y recóndita», escribe, y dice en su propia filosofía: «Yo... lo consideraría una fuerte presunción en su contra, que fuese tan fácil y tan obvia»[86]. Pero este pasaje parece referirse solamente a las discusiones filosóficas o metafísicas en el sentido estricto de la palabra. Porque en lo que sigue inmediantamente, Hume afirma claramente su esperanza de que su propia filosofía de la naturaleza humana pueda proporcionar un atajo hacia todo conocimiento (legítimo): su esperanza de que nos abra el libro de la naturaleza. «Es evidente», escribe, «que todas las ciencias están en relación... con la naturaleza humana... Incluso *la Matemática, la Filosofía Natural y la Religión Natural*, dependen en cierta medida de la ciencia del HOMBRE... Aquí, pues, está el único expediente del que podemos esperar éxito... abandonar el tedioso y dilatado método que hemos seguido hasta hoy, y, en lugar de tomar ahora un castillo y más tarde una aldea en la frontera, *marchar directamente hasta la capital o centro de esas ciencias, a la naturaleza humana misma*; que una vez que la dominemos a ella, *podemos esperar una victoria fácil en todos los demás campos*. Desde esta posición podemos extender nuestras conquistas por todas esas ciencias...»[87].

Este pasaje parece indicar que el propio Hume pensó que la verdad se haría manifiesta una vez que hubiéramos exorcizado al error con ayuda del principio délfico y socrático: «Conócete a ti mismo» (aunque ahora convertido en: Sabe que no puedes «tener... idea de [tu] yo»; porque, como dice Hume, «¿de qué impresión podría derivarse esta idea?»)[88].

---

[86] *Treatise*, Introducción. (Selby-Bigge, pp. xviii s.)
[87] *Treatise*, *loc. cit.* (pp. xix s.) El subrayado es mío.
[88] *Treatise*, Libro I, Parte IV, sección VI. (Selby-Bigge, p. 251.)

El ataque de Hume a las cosas materiales es parte del aspecto metafísico, por así decirlo, de su ataque a las leyes universales, mientras que su negación de la validez de la inducción representa el aspecto lógico de ese ataque. Yo no pienso que sus argumentos metafísicos y epistemológicos sean tan fuertes como sus argumentos lógicos. De hecho no son, ni de lejos, tan fuertes como los argumentos epistemológicos y lingüísticos de Berkeley contra el realismo (o materialismo) de los científicos. Por tanto, me vuelvo ahora hacia Berkeley y hacia sus sucesores modernos.

11. EL STATUS DE LAS TEORIAS Y DE LOS CONCEPTOS TEORICOS

La teoría de Hume, y la de los monistas neutrales, es que el mundo no consiste más que en percepciones y en sus reflejos y que aquéllas están atadas formando mentes y cuerpos físicos. Pero ¿qué ata los manojos?

Según Hume, la ley de la asociación liga las percepciones, formando mentes; y, en la medida en que la causalidad física no es más que hábito, también las liga formando cuerpos. En este sentido, pues, hay una especie de realidad fuera de las ideas: a saber, la ley de su asociación o su tendencia a asociarse.

Según el monismo «neutral», las leyes de la psicología y de la física ligan las ideas. Aquéllas, a su vez, son de dos clases. Existen, en primer lugar, las leyes de la naturaleza, que se obtienen por medio de la observación de los fenómenos y por medio de la inducción por simple generalización a través del hábito o repetición. Y existen, en segundo lugar, teorías científicas más abstractas, especialmente las teorías físicas de Newton. ¿Cuál es su status? ¿Explican nuestras observaciones? Tanto William James como Ernst Mach adoptaron una respuesta berkeleyana a esta pregunta.

Según la concepción de Berkeley, las teorías científicas *no son más que instrumentos* para el cálculo y la predicción de fenómenos inminentes. No describen el mundo ni ningún aspecto de él. No pueden hacerlo porque están *completamente faltas de sentido*. La teoría de Newton no significa nada, porque palabras tales como «fuerza», «gravedad» y «atracción» no significan nada: son *conceptos ocultos*. La de Newton no es una teoría explicativa, sino sólo la ficción de un matemático, el truco de un matemático. Como no describe nada, no puede ser verdadera ni falsa, sólo puede ser útil o inútil, según sirva o no

a su fin predictivo. Berkeley usa el término «hipótesis matemática», para este tipo de truco sin sentido pero útil, ideado en provecho del matemático: «Fabricada y adoptada para abreviar y facilitar los cálculos» como lo expresa Bacon en su crítica de la teoría de Copérnico[89].

Yo estoy en total desacuerdo con *la concepción instrumentalista de las teorías científicas* de Berkeley, y también con su teoría del significado: relega a la categoría de símbolos sin sentido a palabras como «materia» o «sustancia», «corpúsculo» o «átomo» y además a casi todas las palabras empleadas en las teorías de Copérnico, Galileo, Kepler y Newton. Pero nadie antes de Berkeley, y casi nadie después de él, ha estado tan cerca de ver que, según una teoría observacionista o fenomenalista del significado, *todos* los conceptos científicos tienen que carecer por completo de sentido. Porque *todos* los conceptos científicos son «ocultos» en el sentido de Berkeley. No denotan percepciones ni observaciones ni fenómenos, ni tampoco pueden definirse con ayuda de percepciones y observaciones o fenómenos; ni con la ayuda de conceptos que denoten percepciones u observaciones o fenómenos. Es decir, no pueden ser «constituidos», como lo expresé en *L. I. C.*[90] (tampoco pueden ser «definidos operacionalmente»)[91].

Este es el verdadero gran descubrimiento de Berkeley, que él interpretó como una prueba de la carencia de significado de los términos y las teorías científicas. *Yo lo interpreto como una refutación de la teoría observacionista o fenomenalista del significado.*

De hecho, en el sentido de «oculto» de Berkeley, todos nuestros *conceptos* científicos son ocultos; se usan para describir las propiedades estructurales no vistas y, de hecho, invisi-

---

[89] *Novum Organum* II, 36. Respecto a Berkeley, véase mi «Nota sobre Berkeley como precursor de Mach», *El desarrollo del conocimiento científico (C. & R.)*, capítulo 6, donde se dan todas las referencias necesarias.

[90] Véase *L. I. C.*, fin de la sección 25. «Constitución» es el nombre que da Carnap (en *Der logische Aufbau der Welt)* a una definición en términos fenoménicos. Véase también la sección 20, más abajo.

[91] Las definiciones operacionales de los conceptos científicos son tan imposibles como las «constituciones». Tomemos por ejemplo la definición operacional de *longitud* esbozada por Bridgman en su admirable libro *The Logic of Modern Physics.* La longitud se define con ayuda de una descripción de mediciones que incluyen el metro de París (por ejemplo). Ahora bien, como puede verse por el propio análisis de Bridgman, en esta descripción hay implicadas correcciones en razón de la *temperatura.* Pero estas presuponen que hayamos definido *mediciones de temperatura* que, a su vez, presuponen *longitud.* Los círculos de este tipo son esencialmente inevitables. (Véase también la primera nota, y el texto, de la sección VI de «Tres concepciones sobre el conocimiento humano», en *El desarrollo del conocimiento científico* [*C. & R.*], cap. 3 y nota 26 al Apéndice *10 de *L. I. C.)*

bles, de un mundo no visto e invisible[92]. Pero esto no significa que las *teorías* formuladas con la ayuda de esos conceptos sean «ocultas» o «metafísicas» o «no empíricas»: pueden muy bien tener consecuencias contrastables. No me asustan los conceptos científicos «ocultos» precisamente porque tengo un criterio de demarcación que me permite detectar las teorías no científicas por medios que son más efectivos que cualquiera de los supuestos criterios sobre el carácter empírico o no empírico de los conceptos o las palabras implicadas[93].

Volviendo ahora a criticar la interpretación instrumentalista de las teorías científicas de Berkeley, puedo quizá empezar diciendo que no fue lo bastante lejos en su crítica de los conceptos teóricos. Aunque se percató de la inexistencia de lo que él llamaba «ideas generales abastractas», no se percató de la mayoría de los predicados del lenguaje ordinario son también abstractos en ese sentido. El propio Berkeley, cuando *usa* el lenguaje, usa constantemente términos abstractos. Así, cuando habla de «infidelidad» (en el sentido de incredulidad), como, por ejemplo, en esa última embestida de *The Analyst* contra «el moderno aumento de la infidelidad», está claro que no pretende designar con la palabra «infiel» simplemente a un conjunto creciente de semejantes suyos que no le gustan (y que, digamos, se comportan de una cierta manera observable), sino que la usa como un nombre de una *propiedad* verdaderamente *abstracta* de esos hombres. Los términos universales en el lenguaje ordinario no son meramente nombres generales, compartidos (como podría pensarse a la vista de palabras como «mesa» o «perro») por un conjunto concreto de cosas concretas, sino que denotan, como «cobre», «cristal», «manzana», «alimento», «veneno» o «dinero», propiedades estructurales, relacionales, «disposicionales» de las cosas, propiedades que

[92] No creo en espíritus (me refiero a los objetos estudiados por la llamada «investigación psíquica»); no porque estén ocultos, sino más bien porque no están suficientemente ocultos. Son de una especie de ocultación primitiva y representen un ingenuo compromiso entre el mundo ordinario, visible, y la realidad escondida e invisible que el mundo de la ciencia trata de explorar.

[93] Cuando escribí *L. I. C.* —me refiero a la sección 4; véase la nota 1 a dicha sección— pensé, con optimismo, que los positivistas se habían percatado del hecho de que son los enunciados los que son importantes, y no los conceptos, y que habían renunciado al intento de deslindar la metafísica de la ciencia sobre la base del carácter de los conceptos o las palabras que usaba cada una. Pero me equivoqué. El análisis operacional del significado de los conceptos sigue teniendo mucha influencia. E incluso los miembros de lo que fue el Círculo de Viena continúan en la línea de «los antiguos positivistas», como les llamé en 1934. Por ejemplo, Carnap, en su artículo «The Methodological Character of Theoretical Concepts» (en Herbert Feigl y Michael Scriven, eds., *Minnesota Studies in the Philosophy of Science*, vol I, 1951) basa su definición de cláusulas significativas (p. 60) en la de conceptos significativos (p. 51).

son abastractas precisamente en el sentido de Berkeley. Todos los conceptos universales incorporan teorías. Y aunque algunas de ellas puedan ser contrastadas, nunca pueden serlo exhaustivamente (y no pueden ser verificadas nunca). El enunciado «aquí hay un vaso de agua» está abierto a un número indefinido e inexhaustible de contrastaciones —químicas, por ejemplo— porque el agua, como todo lo demás, puede conocerse por su *comportamiento legiforme* (*cf.* el final de la sección 25 de *L. I. C.*). Por tanto, «agua» es disposicional, como todo concepto universal. Incluso «rojo» es disposicional, porque «esta superficie es roja» afirma que esta superficie tiene disposición para reflejar la luz roja. (El hecho de que no existe nada parecido a los universales no disposicionales invalida, por cierto, el programa propuesto por Carnap, en *Testability and Meaning*, de «reducir» las disposiciones a propiedades observables no disposicionales)[94]. Por consiguiente, no sólo la dinámica de Newton, sino la mayoría de los enunciados del lenguaje corriente tendrían que describirse como carentes de significado en el sentido de Berkeley, puesto que «cristal» y, quizá incluso de modo más claro, «agua», a pesar de pertenercer al lenguaje ordinario, son auténticos universales y, por tanto, términos abstractos en el sentido de Berkeley.

De hecho, los términos universales más corrientes del lenguaje ordinario incorporan gran número de teorías empíricas, así como metafísicas o religiosas. Ejemplos llamativos son los términos «padre» e «hijo», que incorporan una teoría empírica aparentemente desconocida en ciertas culturas primitivas, por no hablar de la teoría metafísica de la causalidad que incorporan también. Y, en efecto, la cualidad de «ser un padre» o de «ser un hijo» no puede percibirse, en el sentido de Berkeley, mejor de lo que puede percibirse «ser atraído gravitatoriamente por el Sol». (Una vez más, encontramos que el idealismo de Berkeley choca con su cristianismo. La razón es la misma que antes: la doctrina de la Encarnación es esencial en el cristianismo) [95].

Todo esto indica que la diferencia entre la dinámica de

[94] Véase mi artículo «La demarcación entre la ciencia y la metafísica» (contribución presentada en enero de 1955 al volumen sobre Carnap en la *Library of Living Philosophers*, editada por P. A. Schilpp), especialmente sección 4, desde nota 59 hasta el final de la sección, reimpreso en *El desarrollo del conocimiento científico (C. & R.)*, cap. 11. Me alegra que Carnap haya aceptado ahora esta concepción; porque en la p. 65 del artículo al que hace referencia la nota anterior escribe: «No hay una línea realmente clara entre las propiedades observables y las disposiciones contrastables». Véanse también las dos últimas páginas de mi «Tres concepciones sobre el conocimiento humano», *op. cit.*

[95] *Cf.* sección 7, más arriba, texto a la nota 57.

Newton, por un lado, y los enunciados corrientes del lenguaje corriente, por otro —si es que existe tal diferencia— es sólo una cuestión de grados. Pero esto significa el fin de la teoría de Berkeley. Primero, porque el sentido o la falta de sentido de las palabras difícilmente pueden ser cuestión de grados. (Si una palabra tiene sentido, por poco se sea, tiene sentido.) En segundo lugar, porque es parte de la teoría de Berkeley el que los enunciados del lenguaje ordinario, incluidas las leyes de la naturaleza, son verdaderos o falsos porque describen algo; de donde debemos concluir ahora que lo mismo debe ser cierto de las teorías científicas, tales como la de Newton. En otras palabras, las teorías científicas no son *sólo* instrumentos, sino auténticos enunciados descriptivos. *Son auténticas conjeturas sobre el mundo.*

Esta, naturalmente, es una postura realista: nos hace posible decir que una teoría científica, o una ley de la naturaleza, puede ser verdadera (aunque nunca podamos estar seguros de ello). Y aún más, nos permite decir —como lo hicimos en la sección 5 y, sobre todo, en la 6— que la doctrina metafísica que afirma la existencia de una ley verdadera de la naturaleza está, de uno u otro modo, respaldada por argumentos empíricos.

He tratado de mostrar que un realista no tiene que temer el ataque de Hume a la realidad de las cosas materiales. Pero la debilidad de los argumentos idealistas de Hume no establece la verdad del realismo. El ataque de Berkeley a la interpretación realista de las teorías científicas me parece más interesante y más sutil que su propio ataque, y el de Hume, a las cosas materiales (aunque, personalmente, me siento más atraído por la manera de Hume de buscar la verdad sobre el mundo que por los argumentos lingüísticos de Berkeley): puede aprenderse mucho de este ataque de Berkeley. Y aunque puede ser rechazado fácilmente, mostrando que los argumentos de Berkeley demostrarían demasiado, esto no refuta su interpretación de las teorías científicas como instrumentos. El instrumentalismo puede muy bien ser verdadero, incluso aunque los ingeniosos argumentos de Berkeley en su favor sean insostenibles. Y puesto que una cierta forma de instrumentalismo se ha convertido en un nuevo tipo de ortodoxia desde 1927 (véase el tomo III del *Post Scriptum*), será necesario criticar el instrumentalismo en cuanto tal, independientemente de los argumentos de Berkeley. Para hacerlo podemos dejar ahora nuestra discusión crítica de la teoría subjetivista del conocimiento y volver a nuestro tema principal: la teoría de la ciencia.

12. CRITICA DEL INSTRUMENTALISMO. EL INSTRUMENTALISMO Y EL PROBLEMA DE LA INDUCCION

Aunque creo que he conseguido darle la vuelta a Berkeley y transformar su prueba del instrumentalismo en una prueba en contra de su teoría del significado, hasta el momento no he refutado la doctrina del instrumentalismo como tal; porque puede haber otras versiones diferentes de la de Berkeley.

Por instrumentalismo entiendo la doctrina de que una teoría científica, tal como la de Newton o Einstein o Schrödinger, debe interpretarse como un instrumento, *y nada más que un instrumento*, para la deducción de predicciones de sucesos futuros (especialmente mediciones) y para otras aplicaciones prácticas; y más en especial, que una teoría científica no debe interpretarse como una auténtica conjetura sobre la estructura del mundo o como un auténtico intento de describir ciertos aspectos de nuestro mundo. La doctrina instrumentalista supone que las teorías científicas pueden ser más o menos útiles y más o menos eficaces; pero niega que puedan ser verdaderas o falsas, como los enunciados descriptivos.

Toda la cuestión del instrumentalismo gira en torno a las palabras «nada más que». Porque nadie que sostenga que las teorías científicas son auténticas conjeturas sobre el mundo disputará que pueden ser consideradas *también* como instrumentos para la deducción de predicciones y para otras aplicaciones.

Ya he discutido en *L. I. C.* una cierta forma de instrumentalismo, aunque no bajo ese nombre. Es una concepción discutida por Schlick y que él atribuye a Wittgenstein (véase *L. I. C.*, sección 4, notas 4 y 7 y texto). Esta forma de instrumentalismo es similar a la de Berkeley en la medida en que está de acuerdo con la doctrina de que las teorías científicas carecen de sentido. (Según Schlick, carecen de sentido porque no son «verificables». Una doctrina relacionada de algún modo con el instrumentalismo es el convencionalismo de Poincaré y Duhem que ve en las teorías científicas unas convenciones útiles más que conjeturas que deben ser contrastadas por la experiencia. Como también critiqué esta doctrina con cierta extensión en *L. I. C.*, no me propongo entrar en ella en este capítulo. Pero mi discusión actual, con ciertas alteraciones, podría aplicársele a ella.)

He esbozado brevemente una crítica del instrumentalismo en otro lugar[96]. Pero como el tema es importante en el contexto actual, voy a desarrollar más plenamente esta crítica.

---

[96] *Cf.* mi artículo «Tres concepciones sobre el conocimiento humano», *El desarrollo del conocimiento científico (C. & R.)*, capítulo 5, sección V. (Esta

La idea de mi crítica es ésta: En las ciencias aplicadas y en la tecnología se hace mucho uso de «*reglas de computación*» (como propongo llamarlas). Ejemplos de tales reglas de computación son las reglas y las tablas que se usan para calcular los tiempos de exposición en fotografía (para quienes no tengan un fotómetro). Estas reglas de computación, efectivamente, no son *nada más que* instrumentos; es decir, se proponen, se compran y se venden como instrumentos útiles más que como descripciones informativas del mundo.

Ahora bien, si el instrumentalismo fuese cierto, entonces todas las teorías científicas no serían nada más que reglas de computación. En consecuencia, no podría haber diferencias fundamentales entre las teorías de las llamadas ciencias puras, tales como la dinámica de Newton, y aquéllas reglas tecnológicas de computación que encontramos en todas las ciencias aplicadas y en la ingeniería.

Pero no ocurre así: *existen profundas diferencias entre teorías y meras reglas tecnológicas de computación.* Mientras que el instrumentalismo puede ofrecer una descripción perfecta de esas reglas de computación es totalmente incapaz de dar razón de esas diferencias. Podrían rebatirse mis argumentos estableciendo algo parecido a una regla de traducción que permitiera al instrumentalista traducir a lenguaje instrumentalista cualquier cosa que el teórico pudiera decir para mostrar el carácter peculiar de su enfoque. Pero tal defensa del instrumentalismo no tendría éxito. Porque ese método vaciaría de todo contenido la tesis instrumentalista y la privaría así del interés que tiene en la actualidad. Además hay realmente una diferencia de actitud; por ejemplo, el instrumentalista supone que la gama de sus objetivos prácticos últimos está definitivamente circunscrita. No puede aceptar muy bien la concepción de que es posible que todo pueda servir algún día para algún propósito sin vaciar su tesis de todo contenido. El teórico, por su parte, puede encontrar una teoría interesante en sí misma con total independencia de cualquier consideración sobre su utilidad futura.

Mencionaré aquí diez puntos que arrojan algo de luz sobre las diferencias entre los enfoques del teórico y del instrumentalista y entre teorías y reglas de computación.

---

sección contiene un esbozo de un artículo que aún no se ha publicado, «A Defence of Free Thinking in Quantum Theory», basado en una conferencia que di en Cambridge en 1953.) [Véase el tomo III del *Post Scriptum* y también los siguientes artículos de Popper: «The Propensity Interpretation of Probability», en *The British Journal for the Philosophy of Science* 10, 1959, pp. 25-42; «The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory», en S. Körner, ed., *Observation and Interpretation in the Philosophy of Physics. The Colston Papers*, Vol IX, 1957, pp. 65-70. *Ed.*]

1) La estructura lógica de las teorías es diferente de la de las reglas de computación. Las relaciones lógicas entre dos o más sistemas teóricos son diferentes de las que hay entre dos o más sistemas de reglas de computación. Y las relaciones entre teorías, por un lado, y reglas de computación, por otro, no son simétricas.

Sin discutir a fondo este punto puede observarse que las teorías son sistemas deductivos, de los que se afirma conjeturalmente que son válidos en todo lugar y en todo tiempo. Las reglas de computación pueden presentarse en forma de tabla y se realizan teniendo en mente un fin práctico limitado. Así pues, las normas de navegación útiles para barcos lentos pueden ser inútiles para los veloces aviones. Las reglas de computación (o de navegación, por ejemplo), pueden estar basadas en una teoría (la de Newton, por ejemplo), en un sentido en el que ninguna teoría está basada en reglas de computación.

Este argumento, sin embargo, no es todavía decisivo: podría decirse aún que las propias teorías son algo parecido a reglas de computación glorificadas, super-reglas que se usan para extraer más reglas de computación especiales.

2) Las reglas de computación se escogen sólo por su utilidad. Puede descubrirse que una teoría es falsa, pero puede seguir siendo útil para fines de computación. Por ejemplo, podemos suponer que la teoría de Newton, o la conjunción de la teoría de Newton con la teoría de Maxwell (de Herz) de las hondas sin hilos está falsada. Sin embargo, no hay razón alguna por la cual las reglas de computación utilizadas en navegación (incluido el radar de navegación) no deban seguir basándose en ambas.

3) Al contrastar las teorías, debemos intentar falsarlas. Al poner a prueba nuestros instrumentos, sólo tenemos que saber los límites dentro de los que son aplicables. Si falsamos una teoría, siempre buscamos una mejor. Pero no se rechaza un instrumento porque sea aplicable sólo en casos limitados: contamos de antemano con encontrar esos límites. Si el armazón de un avión es «contrastado hasta su destrucción», después no rechazamos el tipo porque hayamos conseguido destruirlo; podemos rechazarlo porque los límites dentro de los que era aplicable eran más estrechos que los de otro armazón, pero no simplemente porque tuviera esos límites (mientras que sí podemos rechazar una teoría universal simplemente por esa razón).

4) Las aplicaciones de una teoría pueden considerarse como contrastaciones y el que no produzcan los resultados esperados puede llevar al rechazo de la teoría. Esto no ocurre, en la mayor parte de los casos, con las reglas de computación y con otros instrumentos: cualquier fracaso (por ejemplo, el de

las normas tradicionales de navegación en la navegación aérea) puede llevarnos a rechazar la conjetura teórica de que pudieran ser aplicables en un cierto campo, pero continuarán usándose en otros campos (la carretilla puede continuar junto al tractor).

5) Hay una decidida tendencia hacia teorías cada vez más generales por un lado, y hacia instrumentos cada vez más especializados (incluidos los ordenadores), por otro. La segunda tendencia es explicable en términos del instrumentalismo: desde el punto de vista práctico, queremos unos instrumentos que sean los más convenientes para el fin específico que nos ocupa. Así, parece que los intereses y los fines del teórico pueden diferir de los del técnico de ordenadores.

6) Un caso interesante sería el de que tuviéramos ante nosotros dos teorías que fuesen, de momento, indistinguibles en lo que se refiere a sus aplicaciones prácticas previsibles. Porque los instrumentos y las reglas de computación se conciben para un determinado campo de aplicación. Así, el instrumentalista dirá que dos teorías son igualmente útiles si, siendo iguales en otros aspectos, como el de «facilidad de aplicación», llevan al mismo resultado dentro de ese campo. El teórico piensa de modo diferente: si las dos teorías son diferentes lógicamente, intentará *encontrar* un campo de aplicación en el que produzcan resultados diferentes, incluso aunque nadie haya estado interesado antes por este campo. Lo hará porque puede obtener así un experimento crucial que false una de las dos teorías. De este modo, las teorías pueden abrir nuevos campos de experiencia, campos en los que nadie había pensado anteriormente. A esta función exploratoria (e intelectual) de las teorías puede quitársele importancia diciendo que las teorías son instrumentos para la exploración (como el barco de un explorador o el microscopio). *Pero entonces, tendríamos que admitir que hay una realidad por explorar; y en ese caso, puede ser descrita, verdadera o falsamente,* que es precisamente lo que el instrumentalismo berkeleyano quiere negar. (Digo «instrumentalismo berkeleyano» porque está claro que el instrumentalismo puede combinarse también con un realismo metafísico, e incluso con el materialismo. Por ejemplo, podemos negar que las teorías describen la realidad, pero afirmar que hay cosas en sí mismas —incluso cosas materiales en sí mismas— que, aunque quizá no podamos describirlas nunca por medio de nuestras teorías, sí podemos operar sobre ellas, con mayor o menor éxito, por medio de nuestros instrumentos, incluidas nuestras teorías.)

7) Esto me lleva a una importante distinción entre dos tipos de predicciones (o dos sentidos del término «predicción») cuya diferencia no puede ser apreciada, aparentemente, por el instrumentalista. Uno de los tipos es, por ejemplo, la predicción

del próximo eclipse, la hora en que ocurrirá, la región de visibilidad, etc., o la predicción del número de guisantes de diferente color en experimentos mendelianos sobre la reproducción: hablando en términos generales, la predicción de sucesos de una clase conocida. El otro tipo es la predicción de un suceso de una clase que nunca se había contemplado seriamente antes de formular la nueva teoría; un suceso de cuya posibilidad nos enteramos, por así decirlo, a través de la teoría. Ejemplos de esta segunda clase son las predicciones de Einstein respecto a los eclipses solares; o sus predicciones respecto a la conversión de la masa en energía. Las prediciones de este tipo surgen a menudo de una nueva teoría para sorpresa de su autor, cuya única intención puede muy bien haber sido eliminar algunas de las dificultades de la teoría vigente en este momento. Abren un nuevo e inesperado mundo de hechos, o quizá un nuevo aspecto de nuestro mundo; y sería difícil encajar sin esfuerzo esta situación dentro de un marco de referencia instrumentalista.

8) La cuestión crucial es, desde luego, si la teoría tiene o no, más allá de su capacidad como instrumento, algún contenido informativo. En vista del punto anterior, esto difícilmente puede negarse, porque si por la teoría podemos aprender algo sobre sucesos de una clase desconocida, la teoría debe ser capaz de *descubrirnos* esos sucesos (como de hecho hace). Pero hay argumentos más directos. ¿Quién negará hoy que la teoría copernicana describe la estructura (aproximada) de nuestro sistema solar? Sin embargo, por sugerencia de Osiander, el editor de Copérnico, que propuso una interpretación instrumentalista del sistema copernicano, lo negaron el Cardenal Bellarmino, Francis Bacon y el Obispo Berkeley. Lo que esos hombres negaban era, precisamente, que el sistema copernicano tuviera contenido descriptivo o informativo.

9) Las teorías, al contrario que las reglas de computación, nos ayudan a *interpretar* nuestras experiencias. Un ejemplo es la interpretación, por Lise Meitner y Otto Frisch, de las observaciones de Hahn como observaciones de la fisión del uranio. Naturalmente, no es ningún accidente que Frisch fuese uno de los autores de la teoría del núcleo como gota que predecía que los núcleos muy pesados se dividirían. (Este hecho puede explicar también por qué Frisch descuidó el problema de la emisión y recaptura de neutrones: no formaba parte de la teoría del núcleo como gota.) Pero lo que ocurrió aquí, en la región fronteriza de la ciencia, ocurre continuamente en la vida corriente. Constantemente interpretamos nuestra experiencia con ayuda de teorías. Un mal sabor o mal olor se interpretan como causados por un huevo podrido; y un patinazo —un término com-

pletamente teórico— *explica* o interpreta un movimiento inusual y peligroso de un coche como debido a una insuficiente fricción entre los neumáticos y la carretera. (Este es un ejemplo del modo en que la línea divisoria —si existe tal cosa— entre el lenguaje teórico y el lenguaje ordinario está constantemente cambiando.)

10) La predicción ha sido considerada por los instrumentalistas, desde Berkeley hasta nuestros días, como una de las principales tareas de la ciencia. El valor práctico de las predicciones es obvio y no necesita que el instrumentalismo lo justifique. Los teóricos, sin embargo, pueden explicar la importancia de las prediciones para la ciencia en sus propios términos, sin suponer que la predicción es una tarea impuesta a la ciencia desde fuera. Para el teórico, las predicciones son importantes casi exclusivamente por su relación con la teoría; porque él está interesado en la búsqueda de teorías *verdaderas* y porque *las predicciones pueden servir como contrastaciones* y proporcionar una oportunidad para la eliminación de teorías falsas.

La segunda clase de predicciones mencionada más arriba —la predicción de un suceso insospechado hasta el momento— nos da, no sólo cierta medida de la novedad de una teoría, sino también la medida de su superioridad sobre la teoría antigua y, por ello, del avance realizado. Incluso aunque la nueva teoría sea rechazada algún día, puede habernos llevado a descubrir nuevos tipos de hechos empíricos.

Sin duda, los diez puntos pueden ser disputados por un instrumentalista decidido. Pero a mí me parece que justifican ampliamente el rechazo del instrumentalismo.

Para cualquiera que adopte una concepción instrumentalista, *desaparece el problema de la inducción.* No existe la cuestión de *la verdad o falsedad* de los instrumentos. Por consiguiente, no existe la cuestión de la *validez* de los procedimientos o técnicas usadas en la creación o el perfeccionamiento de los instrumentos. Y el problema de la inducción sólo se ocupa de cuestiones de verdad, falsedad y validez.

El instrumentalismo debe, sin duda, mucho al deseo de resolver, o disolver, el problema de la inducción. Berkeley creía en la inducción de generalizaciones simples; por ejemplo, que nubes de cierto tipo normalmente traen lluvia, pero veía que una teoría «oculta» —tal como la de Newton o el atomismo— no podía ser el resultado de una inducción. Esto no le creaba ninguna dificultad, sin embargo, puesto que sostenía que las teorías «ocultas», si tenían éxito, no eran nada más que instrumentos. La solución de Berkeley al problema estaba casi olvidada pero fue redescubierta por Mach, que enseñaba que las teorías no eran más que resúmenes económicos e instrumentos

de predicción; por Poincaré y Duhem, que enseñaron que no eran más que convenciones; y por J. S. Mill, Wittgenstein y Schlick (y más tarde por Ryle), que dijeron que no eran auténticas proposiciones, sino pseudo-proposiciones cuya función era servir como reglas de inferencia (o «licencias de inferencia»), como reglas para transformar auténticos enunciados de observación en otros enunciados auténticos de observación (es decir, en predicciones).

Berkeley, Mach, Poincaré y Duhem creían, todos ellos, que las generalizaciones simples de un orden inferior de universalidad se basaban en la inducción. Pero veían, al menos, que las teorías más abastractas no podían basarse en la inducción. Hume fue más lejos y señaló que la inducción incluso de generalizaciones simples, no era válida y era imposible de justificar; pero nunca dudó de lo que consideraba un hecho de la psicología o de la biología: que los hombres y los animales (influenciados por la asociación o por el hábito) realizan inducciones que tienen mucho éxito, por inválido que sea todo intento de justificarlas.

Yo voy más lejos que Hume: sostengo que los procedimientos inductivos simplemente no existen (ni siquiera los de nivel inferior) y que la historia de su existencia es un mito[97]. Berkeley, Poincaré y Duhem tenían razón al enseñar que es imposible obtener una teoría abstracta o de nivel superior, tal como la de Newton, por inducción a partir de la observación. Pero estaban equivocados al pensar que hay una diferencia esencial, en este aspecto, entre leyes de nivel superior e inferior; porque todas las leyes son invenciones útiles más que generalizaciones inductivas. (Mi propia concepción, naturalmente, es que no son meramente útiles, sino auténticas conjeturas sobre la estructura del mundo que pueden ser constrastadas.) Sin duda, las hipótesis de nivel inferior no son tan abstractas como las teorías explicativas de nivel superior. Pero son, no obstante, teóricas y abstractras.

Volvamos ahora a la sugerencia de Mill, Wittgenstein y Schlick de que debemos considerar las leyes o teorías universales como reglas de inferencia en vez de como enunciados genuinos[98].

---

[97] Véanse las secciones 3 y 9 más arriba y mi artículo «Philosophy of Science: A Personal Report», especialmente las secciones IV a VI, reimpreso como capítulo 1 de *El desarrollo del conocimiento científico (C. & R.)*.

[98] Se recordará que esta sugerencia la hicieron Wittgenstein y Schlick para reconciliar la inverificabilidad de las leyes universales con el criterio de sentido del *Tractatus;* véanse las notas *4, 7 y 8 a la sección 4 y la nota 1 al Apéndice *1 de *L. I. C.* La referencia a Mill alude a su *Logic,* Libro II, capítulo III, 3: «Toda inferencia es de particulares a particulares». Véase también la primera

La sugerencia puede desarrollarse como sigue: tomemos un enunciado universal del tipo «todos los hombres son bípedos». En su presencia —es decir, si lo usamos tácitamente como premisa, o como premisa «suprimida»— podemos inferir, del enunciado singular «Sócrates es un hombre» la conclusión «Sócrates es un bípedo». Supongamos ahora que sólo los enunciados singulares tienen sentido; o, más bien, tomemos el supuesto más débil de que, por una u otra razón, nos interesan sólo los enunciados singulares (digamos que porque sólo los enunciados singulares describen hechos empíricos) y no los enunciados universales. Puede seguir interesándonos, pues, la transición, o inferencia, de «Sócrates es un hombre» a «Sócrates es un bípedo»; y podemos decir que todo el interés del enunciado universal «todos los hombres son bípedos» reside en el hecho de que nos permite efectuar la transición desde un enunciado singular a otro. Dicho de otro modo, podemos interpetar que el enunciado universal nos da derecho a efectuar esta transición o que *da validez a ciertas inferencias*, más que como una aserción sobre un «hecho universal» cualquiera que sea el significado que eso tenga.

Ciertamente podemos considerar los enunciados universales como validadores de ciertas inferencias, o como equivalentes a ciertas reglas de inferencia: es un simple hecho de la lógica que, siempre que tengamos una inferencia válida que proceda de más de una premisa, podemos interpretar cualquiera de las premisas (si son verdaderas) como una regla (válida) que nos permite sacar la conclusión a partir de las otras premisas. Así pues, podemos decir que si «Sócrates es un hombre» es verdadera, valida la inferencia de «todos los hombres son bípedos» a «Sócrates es un bípedo»; y del mismo modo, que si «todos los hombres son bípedos» es verdadera, valida la inferencia de «*x* es un hombre» a «*x* es un bípedo». Sin embargo, no puede interpretarse que este hecho establece que un enunciado universal es, esencialmente, una regla de inferencia o nada más que una regla de inferencia.

Para discutir esta cuestión, convengamos en decir de los enunciados que pueden ser *verdaderos* o *falsos*, y de las inferencias, o reglas de inferencia, que pueden ser *válidas* o *no válidas*. (Decimos que son válidas si mantienen la verdad; es decir, si son tales que la verdad de las premisas asegura la verdad de la conclusión.) Podemos decir entonces que en los lenguajes usuales, incluido el lenguaje matemático, el de la ciencia y el inglés corriente, un enunciado universal como «todos los hombres son

---

nota a la sección IV de mi artículo «Tres concepciones sobre el conocimiento humano», *op. cit.*

bípedos» es verdadero de hecho si, y sólo si, la transición de «*x* es un hombre» a «*x* es un bípedo» mantiene *de hecho* la verdad (o, podemos decir, es «válido *de hecho*»). Además, un enunciado universal será *lógicamente* verdadero (o «analítico») si, y sólo si, la inferencia, o la regla de inferencia correspondiente, mantiene *lógicamente* la verdad, o es *lógicamente* válida.

A causa de estas equivalencias, no se gana nada si «explicamos» los enunciados universales como reglas de inferencia, a menos, claro, que queramos hacer la propuesta de *sustituir* los enunciados universales por reglas de inferencia, quizá porque pensemos que las reglas de inferencia puedan tener menos probabilidades que los enunciados universales de crear confusiones o rompecabezas filosóficos. Pero esto equivaldría a hacer la propuesta de *evitar* o *descartar* los enunciados universales y sustituirlos por reglas de inferencia o, dicho de otro modo, a hacer la propuesta de que se use un lenguaje en el cual no haya enunciados universales y estén reemplazados por reglas de inferencia.

Podría construirse sin dificultades un lenguaje de este tipo. No sería un lenguaje corriente, sino uno artificial, en la medida en que su gramática prohibiría el uso de enunciados universales. Todos sus enunciados serían singulares y podrían ser, por tanto (en el sentido del *Tractatus*), funciones de verdad de los enunciados atómicos y podrían ser verificables. Además, al expurgar todos los enunciados universales, podría parecer que desaparecía el problema de la inducción[99].

Un lenguaje artificial como éste se conformaría, sin duda, a muchas de las intenciones positivistas. Sin embargo, ni la posibilidad de construirlo ni su construcción y su adopción reales ayudarían a resolver ningún problema. Tampoco clarificarían

[99] Un modo particular de construir un lenguaje *L* en el que los enunciados universales estén prohibidos (y sustituidos por la debida licencia de inferencia) es éste; establecemos la regla general de que siempre que decidamos aceptar, en los idiomas corrientes, el enunciado «todas las *A*es son *B*s», aceptaremos en *L* todos los enunciados singulares obtenibles por sustitución de «si *x* es una *A*, entonces *x* es una *B*». Esta regla tiene el efecto deseado. Si formulamos esta regla para eliminar enunciados universales como una regla para eliminar sistemas teóricos universales, podemos expresarla de manera siguiente: Sustituya cualquier axioma universal «todas las *A*es son *B*s» de la teoría por el axioma-figura silogística «si *x* es una *A*, entonces *x* es una *B*» (!). O, con otras palabras: Acepte cualquier enunciado singular que pueda obtenerse por sustitución a partir de esta figura silogística como un axioma de la teoría (!).

Esta regla llevará, por lo general, a teorías en las que infinitamente muchos axiomas singulares sustituirán a cada axioma de la teoría ordinaria. La situación es muy análoga a la que se produce al aplicar el teorema de W. Craig a la eliminación de una cierta clase de predicados, por ejemplo a «predicados teóricos»; y la confusión filosófica creada en ambos casos es también análoga (véase también *L. I. C.*, nota 4 al Apéndice *X).

nada. De hecho, la propuesta de sustituir los enunciados universales por reglas de inferencia, al sugerir pseudo-soluciones, ha creado nuevas confusiones y enredos filosóficos que, a juzgar por la experiencia del pasado, tardarán en disiparse.

Debido a la equivalencia de los enunciados universales y las reglas de inferencia, todo lo que puede conseguir la propuesta es sustituir el problema de la verdad de los enunciados universales por el de la validez de las reglas de inferencia correspondientes. No hay ventaja alguna en esta sustitución, puesto que un problema es precisamente equivalente al otro. Así, parece que los filósofos que propusieron esta sustitución deben de haber estado confusos respecto a la cuestión. Creyeron ver una ventaja o una solución donde no había ninguna de las dos cosas.

Pero la situación es peor aún.

Es habitual no usar más que reglas de inferencia *lógicas* o *analíticas*. Con la sugerencia de que todos los enunciados universales se consideren reglas de inferencia, se nos incita implícitamente a tomar todos los enunciados universales como analíticos. (Esto equivaldría a adoptar una forma radical de convencionalismo: véanse los capítulos IV y VII de *L. I. C.*) La costumbre de usar sólo reglas de inferencia lógicas o analíticas está relacionada con el hecho de que en la vida corriente, igual que en las matemáticas o en la ciencia, casi nunca somos conscientes de las reglas de inferencia que usamos. Las usamos inconscientemente, implícitamente, confiando en su validez sin hacer preguntas: difícilmente puede existir algo como una regla de inferencia *problemática*[100]. En consecuencia, raramente ponemos en duda la validez de las reglas de inferencia y nunca pensamos en contrastarlas experimentalmente. Pero si, como se nos propone, fuésemos a interpretar todas las teorías universales como reglas de inferencia, entonces deberíamos tratar esas reglas de inferencia como teorías universales: deberíamos contrastarlas, tratar de falsarlas, a menos que renunciemos al método crítico de la ciencia.

Los filósofos que han defendido esta propuesta no parecen haber percibido estas consecuencias. Al considerar las leyes científicas como analíticas, se acercaron al convencionalismo. Y, sin embargo, entre ellos había algunos destacados oponentes del convencionalismo.

Todo esto puede indicar la confusión que ha creado la propuesta de que los enunciados universales se consideren reglas de inferencia.

---

[100] De la teoría de la probabilidad podemos obtener un sistema de reglas de inferencia casi válidas; con todo, aunque son sólo casi válidas, esas reglas no son en absoluto problemáticas.

Más recientemente se ha puesto en boga una nueva actitud hacia el problema de la inducción. Entre los filósofos se ha extendido la idea de que la inducción puede ser en realidad un mito y que el problema de la inducción puede resolverse en realidad señalando que no existe tal cosa como la inducción. Pero algunos filósofos siguen impresionados por el hecho de que muchos científicos sostengan que ellos hacen inducciones, que obtienen leyes generalizando los resultados de sus observaciones.

De esos dos ingredientes se deriva la siguiente teoría: Los científicos, nos dicen, realizan lo que puede llamarse «inducciones» en diversos campos de la ciencia, en cada uno según su metodología o procedimiento peculiar. Pero esas «inducciones» no deben confundirse con inferencias: hacer inducciones es simplemente una habilidad que se enseña por el ejemplo y se adquiere con la práctica científica. Puesto que no hay problema respecto a la validez de las habilidades, el problema de la inducción —que es el problema de la validez de la inducción— desaparece. En resumen, la *inferencia* inductiva es un mito: sólo existen *procedimientos* inductivos o *técnicas* inductivas (en realidad, hay gran variedad de ambos); pero éstos no pueden dar lugar a ningún «problema de la inducción».

Esta teoría no tiene nada que ofrecer a nadie si no es a un instrumentalista. Y para él, es superflua. Porque si las teorías científicas no son más que instrumentos, entonces el problema de la inducción no surge de ningún modo, como hemos visto.

Aquellos de nosotros que, por el contrario, no aceptamos el instrumentalismo, sino que consideramos que las teorías científicas son hipótesis o conjeturas —los resultados provisionales de nuestra búsqueda de teorías verdaderas— encontraremos que el decir que hay *una multiplicidad de procedimientos inductivos* no es de gran ayuda: no puede hacer más que multiplicar las dificultades del problema de la inducción. En contraposición, el decir que *no hay procedimientos inductivos de ninguna clase* (aunque hay gran riqueza de procedimientos por los que podemos contrastar nuestras teorías, procedimientos que pueden ser minuciosamente analizados en términos de lógica deductiva) puede ayudarnos a eliminar el problema de la inducción y así, a resolverlo.

13. EL INSTRUMENTALISMO CONTRA LA CIENCIA

Según el instrumentalista, las teorías científicas no pueden ser descubrimientos reales: son artefactos. La ciencia es una actividad que consiste en la fabricación de artefactos: fontane-

ría glorificada. (En el mundo de las apariencias, en el mundo sin enigmas, no hay lugar para los auténticos descubrimientos científicos.) La filosofía del instrumentalismo, por otro lado, resulta ser un auténtico descubrimiento científico: revela, desvela, el verdadero carácter de la ciencia, la realidad tras su apariencia. Revela que las teorías científicas no son lo que parecen. No son, como parecen, descripciones de un mundo tras el mundo de la apariencia: el análisis filosófico nos la revela como meros instrumentos. (Y los «grandes científicos» se nos revelan, mediante el análisis filosófico, como fontaneros glorificados. ¡Qué alivio!)

La tendencia del instrumentalista es anti-racionalista. Implica que la razón humana no puede descubrir ningún secreto de nuestro mundo. De modo que no sabemos más sobre el mundo hoy de lo que sabíamos hace cuatrocientos años. Nuestro conocimiento de los hechos no ha aumentado: sólo nuestra habilidad para manejarlos y nuestro conocimiento sobre cómo construir artefactos. No hay revolución científica, según el instrumentalismo: sólo hay una revolución industrial. No hay verdad en la ciencia: sólo hay utilidad. La ciencia es incapaz de iluminar nuestras mentes: sólo puede llenarnos la barriga.

No es sorprendente encontrar estas concepciones en el Obispo Berkeley. De hecho, Berkeley dejó bien claro que las publicaba sobre todo con la esperanza de defender a la religión contra el asalto de la ciencia y del «libre-pensamiento»; contra la pretensión de que la razón, sin la ayuda de la divina revelación, puede descubrir un mundo tras el mundo de la apariencia. Pero es un poco sorprendente encontrar apoyo para esas concepciones instrumentalistas en el campo de los admiradores de la ciencia.

Mi primer ejemplo de un admirador de la ciencia que sostiene el instrumentalismo es Reichenbach. No hay más dudas sobre la lealtad de Reichenbach que sobre la del Obispo: Reichenbach está en el campo contrario. Cree en la enseñanza de la ciencia. «El filósofo debe admitir», escribe, «que la naturaleza puede muy bien ser tal como la describe el físico cuántico; el hecho de si esta descripción es verdadera, es asunto de la ciencia empírica»[101]. Cabe tanta duda sobre el fundamental realismo de Reichenbach como de las convicciones realistas de Hume o de Russell. Además, Reichenbach —como los «monistas neutrales»— se aferra a la creencia de que su análisis de la

[101] La cita es la última oración del artículo de Reichenbach, «The Principle of Anomaly», *Dialéctica* 2, 1948, descrito por el autor como un resumen perfeccionado de su *Philosophical Foundations of Quantum Mechanics*, 1944.

física es filosóficamente neutral. «Los que afirman que la causalidad es *a priori* y se extiende a los inobservables y los que se niegan a hablar del comportamiento causal de los inobservables, se parecen en que cometen el error de juzgar a la mecánica cuántica desde el punto de vista de ciertas doctrinas filosóficas. Es posible hablar sobre el status lógico de la mecánica cuántica en términos neutrales»[102]. Esta posibilidad afirma haberla realizado él en su investigación: «La investigación se lleva a cabo sin la presuposición de ninguna concepción filosófica», escribe[103]. Pero, de hecho, no es más neutral que los monistas neutrales: acepta implícitamente la teoría subjetiva del conocimiento. Es un idealista del estilo de Berkeley o de Hume; y como instrumentalista berkeleyano, está mejor informado sobre los objetos de que trata la física que la propia física. Porque todo su análisis de la física cuántica se basa en la distinción entre los *fenómenos* observados, tales como ese árbol al otro lado de mi ventana mientras lo estoy mirando, y las *interpolaciones* entre esos fenómenos, que él llama «interfenómenos», tales como ese mismo árbol cuando he bajado la vista para escribir esta oración. «Si decimos que un árbol existe mientras no estamos mirándolo», escribe Reichenbach, «...interpolamos un objeto no observado entre observables; y seleccionamos el objeto interpolado de tal modo que nos permite establecer *...el postulado de idéntica causalidad* para objetos observados y no observados...»[104]. Y continúa: «Si abandonamos este postulado, llegamos a diferentes descripciones de los inobservables. Esas descripciones no son falsas... El lenguaje habitual, según el cual el objeto persiste cuando cesa la observación, se construye eligiendo una descripción entre las de esa clase; esta descripción normal, o *sistema normal*, está determinada por el postulado de idéntica causalidad para los objetos observados y no observados. *El postulado no es ni verdadero ni falso, sino una regla que usamos para simplificar nuestro lenguaje*»[105].

Consideremos esta doctrina un momento. Lo que dice Reichenbach —lo que quiere decir, como lo muestran su libro y el artículo— es esto: los objetos que interpolamos entre observaciones están más o menos *arbitrariamente escogidos* o «seleccionados». El «sistema normal» consiste en seleccionar los objetos de tal manera que las leyes causales para los objetos observa-

[102] *Loc. cit.*

[103] *Cf.* el final de su *Summary*, *op. cit.*, p. 349.

[104] *Op. cit.* p. 341. Los puntos suspensivos en esta cita indican una larga omisión de más de tres oraciones. No obstante, esta cita presenta la doctrina de Reichenbach con exactitud.

[105] *Loc. cit.* La última oración de la cita la he subrayado yo.

dos y no observados permanezcan iguales. Pero esta elección «normal» es meramente la opción libre de un postulado que no es ni verdadero ni falso; y lo preferimos sencillamente porque queremos simplificar nuestro lenguaje. (Así pues, el «postulado» puede describirse como una convención.)

Reichenbach usa el ejemplo siguiente: «...podríamos suponer», escribe, «que el árbol no observado se divide en dos árboles»[106]. Pero en este caso, explica, es posible que tengamos que alterar nuestras leyes causales, por lo que respecta a los objetos no observados; porque podemos observar la sombra del árbol, aunque no observemos el árbol; y tendríamos que suponer que los dos árboles (en los que se divide el árbol mientras no lo observamos) «causan una sola sombra»[107].

Tales suposiciones serían claramente extrañas; y esa es la única razón, según Reichenbach, por la que habitualmente adoptamos el «sistema normal», es decir, el «postulados de idéntica causalidad para objetos observados y no observados»: lo adoptamos para «simplificar nuestro lenguaje».

En contra de esta concepción quiero señalar que la división del árbol cuando no estamos mirando y su recomposición cuando volvemos a mirar violaría todas las leyes de la física. Porque implicaría (según la física) fuerzas muy considerables, aceleraciones, desprendimiento de energía, etc. Y esto equivale a decir que *la física nos informa sin ambigüedades de que el árbol no se divide cuando no lo miramos.* (Ni tampoco desaparece, según la física: eso supondría destrucción de materia y, por tanto, algo de la naturaleza de una explosión atómica.) Así pues, cuando Reichenbach dice que la suposición «normal» —según la cual el árbol no observado se comporta como si lo estuvieran observando— «no es ni verdadera ni falsa», sino un simple mecanismo «para simplificar nuestro lenguaje», nos revela que las claras e inequívocas aseveraciones de la física son, en realidad, sólo mecanismo de simplificación. Y cuando afirma que podemos decidir adoptar otros «postulados» (naturalmente, siempre que produzcan los mismos resultados en lo que respecta a los fenómenos u observaciones), entonces está implícitamente de acuerdo con Berkeley y Mach en que las teorías físicas son sólo instrumentos para la descripción y predicción de fenómenos.

Así, Reichenbach es conducido por su teoría subjetivista del conocimiento a una posición que aborrece: la creencia de que el filósofo puede saber más, o saber mejor, que el físico; a la

---

[106] *Loc. cit.*
[107] *Loc. cit.*

creencia de que la filosofía puede corregir las aseveraciones de la física.

Esta posición idealista de Reichenbach resulta ser esencial para su solución a los problemas de la teoría cuántica. Porque su resultado es éste: En la teoría cuántica es «imposible dar una definición de interfenómenos de tal manera que los postulados de causalidad sean satisfechos»[108]. Así, según la teoría cuántica, «*la clase de todas las descripciones de interfenómenos no contiene sistema normal alguno*»[109]. Reichenbach llama a este resultado, del cual pretende haber dado una prueba, «el principio de anomalía» (o, a veces, «el principio de anomalía causal»).

Ahora bien, desde el punto de vista idealista, éste es efectivamente un resultado. Porque puede formularse como sigue: La teoría cuántica cumple lo prometido. Nos permite calcular correctamente las observaciones. Cierto, no nos permite introducir objetos no observados de la manera normal. Pero como la manera normal de introducir objetos no observados es, en cualquier caso, una cuestión de pura conveniencia y de simplicidad lingüística, su resultado negativo no tiene por qué preocuparnos.

Me propongo discutir en el tomo III de este *Post Scriptum* el problema de la interpretación de la teoría cuántica desde mi propio punto de vista realista (sin volver a referirme a Reichenbach, por cierto). Aquí sólo quiero decir que la solución de Reichenbach me parece insatisfactoria, como todas las soluciones idealistas, porque es demasiado simple. A pesar del impresionante aparato matemático y lógico que usa, sufre de la vieja enfermedad: es fatalmente fácil. Porque equivale a decir: «El formalismo cumple lo prometido ¿Qué más quieren ustedes? ¿Quieren entenderlo? Eso es fácil: no es más que un instrumento para la predicción de fenómenos; y entender un instrumento significa saber cómo está construido y cómo funciona. Si quieren saber más —si quieren, por ejemplo, tener una idea de cómo interactúan los átomos o los electrones o los fotones— entonces deben haber olvidado que las entidades no observadas (tales como átomos, electrones o fotones) no son más que interfenómenos, introducidos sólo para conveniencia nuestra; de modo que su pregunta no tiene sentido; incluso puede que estén haciendo una pseudo-pregunta metafísica».

Esto no es, desde luego, una paráfrasis de lo que Reichenbach dice efectivamente, pero tampoco es una parodia. Por el contrario, es una versión bastante ajustada de sus resultados y de su actitud (que es semejante a la de Bohr y Heisenberg). En

[108] *Philosophical Foundations of Quantum Mechanics*, 1944, pp. 32 s.
[109] *Loc. cit.* El subrayado es de Reinchenbach.

cualquier caso, lo más importante que quiero destacar con mi versión es el idealismo subjetivista de Reinchenbach y el hecho de que es un ingrediente esencial de sus resultados y no simplemente un lapsus, o sólo una forma de decir cosas que podrían también haberse expresado en el «lenguaje del realismo».

Mi segundo ejemplo es Carnap. Como en el caso de Reichenbach, nadie puede dudar de lo sincero de la lealtad de Carnap a la ciencia. En sus primeros tiempos, especialmente en *Der logische Aufbau der Welt*, Carnap trató de construir un sistema del mundo de la ciencia sobre los cimientos de la experiencia sensorial subjetiva de una persona; pero poco después, Otto Neurath le convirtió al «fisicalismo». En su largo libro —*Logical Foundations of Probability*— desarrolla una teoría probabilística de la inducción científica. Esta teoría lleva al resultado, correcto en mi opinión (*cf. L. I. C.*, sección 80), de que en un universo infinito, la probabilidad lógica de una ley universal, incluso con la evidencia empírica más favorable, siempre será igual a cero. Pero, desgraciadamente, Carnap identifica el grado de confirmación de una ley con su probabilidad lógica. Obtiene así el resultado, no deseado, de que el grado de confirmación de cualquier ley (sea aceptada o rechazada) debe ser también igual a cero. Así, *ninguna ley es confirmable*. Este resultado le hace plantearse la cuestión (p. 574): «¿Son necesarias las leyes para hacer predicciones?». Su respuesta es que no lo son y, por tanto, que «no son indispensables» para la ciencia. El argumento avanza así un paso más que el instrumentalismo. Como los instrumentalistas, Carnap supone que la única función de la ciencia es hacer predicciones; como ellos, cree que las leyes son «instrumentos eficaces» para este fin; pero sostiene también que no son necesarias. De hecho, su argumento podría mostrar que son completamente superfluas. «Vemos», escribe Carnap, «que el uso de las leyes no es indispensable para hacer predicciones. No obstante, es *conveniente*, naturalmente, enunciar leyes universales en libros de física, biología, psicología, etc. Aunque esas leyes enunciadas por los científicos no tengan un alto grado de confirmación[110], ...sirven como *instrumentos eficaces* para descubrir esas predicciones singulares muy confirmadas que son necesarias en la vida práctica»[111].

[110] Esta es una subestimación de Carnap: en un universo infinito, el grado es cero.

[111] *Logical Foundations of Probability*, 1950, p. 575. Los subrayados son míos. Mis puntos sustituyen las palabras «tienen una confirmación de casos sumamente idónea y por tanto». Las he omitido porque no quiero explicar aquí, ni criticar, el concepto de Carnap de «una confirmación idónea de

Admitamos que las leyes universales son enunciados en la teoría de Carnap; pero como resulta que no son confirmables, ni tampoco indispensables para la ciencia —aunque son «instrumentos eficaces para las predicciones»— está claro que la teoría de la inducción de Carnap le ha llevado a una concepción estrechamente afín a la de Reichenbach; sólo los enunciados singulares de observaciones posibles importan realmente; porque el carácter de las leyes es simplemente el de enunciados no confirmables cuya sola función es que son eficaces pero, de todos modos, instrumentos de predicción *completamente* superfluos.

Considerando la actitud totalmente diferente que adopta Carnap en un lugar anterior del mismo libro (estoy pensando sobre todo en su referencia a Einstein y a mí mismo en la página 193, que muestra que su intención era hacer justicia a las teorías), sin duda, él ni pretendía ni esperaba esos resultados. Son consecuencias no buscadas de su aceptación implícita de una epistemología inductivista[112].

En los volúmenes posteriores de este *Post Scriptum* encontraremos más ejemplos del desafortunado papel representado por las interpretaciones subjetivistas, inductivistas e instrumentalistas de la ciencia y, más especialmente, de la física cuántica.

14. LA CIENCIA CONTRA EL INSTRUMENTALISMO

Podemos concluir nuestra discusión de la teoría subjetiva del conocimiento. La cuestión de la sensatez del realismo metafísico se planteó (al final de la sección 6) en relación con la actitud escéptica de Hume. Nuestro análisis de la teoría subjetiva del conocimiento, de Hume y sus sucesores, ha llevado a la conclusión de que en esta teoría no se encuentran argumentos serios contra lo razonable del realismo metafísico.

---

casos», cuya insuficiencia he señalado en otro lugar; véase *The British Journal for the Philosophy of Science* 6, 1955, pp. 157-163, y 7, 1956, pp. 249-256, especialmente sección 7, pp. 251 ss.

[112] Critiqué las conclusiones de Carnap sobre las leyes universales más extensamente en «La demarcación entre la ciencia y la metafísica», contribución al volumen sobre Carnap de P. A. Schilpp, ed., *Library of Living Philosophers;* reimpreso en *El desarrollo del conocimiento científico (C. & R.),* cap. 11. El artículo posterior de Carnap, «The Methodological Character of Scientific Concepts» (véase nota 93, sección 11, más arriba), revisa algunos de los resultados que yo había criticado. El objetivo de su nuevo artículo es readmitir a las teorías como significativas y dar un criterio de su significación. Creo, sin embargo, que puede mostrarse fácilmente que su nuevo criterio de significación es al mismo tiempo *demasiado estrecho y demasiado amplio,* exactamente como el viejo criterio de sentido positivista, que critiqué, por esta misma razón, en la sección 8 de *L. I. C.* (véase sección 19, más adelante).

La realidad de los cuerpos físicos está implícita en casi todos los enunciados de sentido común que formulamos; y esto, a su vez, entraña la existencia de leyes de la naturaleza: así, todas las aserciones de la ciencia suponen realismo. Estos argumentos hacen razonable el creer que hay leyes verdaderas de la naturaleza, aunque esta concepción no es verificable ni falsable y es, por tanto, metafísica. Nuestra discusión del escepticismo de Hume y del de sus seguidores no ha sacado a la luz nada de peso contra estos argumentos. Sólo encontramos un especioso dogma: que lo que hay en mi mente (o en mi lenguaje) sólo puede reflejar lo que previamente ha estado en mis sentidos. En su camino encontramos incluso una serie de consecuencias no buscadas e inaceptables, incluidas contradicciones, francamente descritas como tales por Hume.

Quizá la cosa más notable que descubrimos fue la actitud de muchos miembros de esta escuela de pensamiento respecto a sus propios resultados. Del mismo modo que Frege que, ante las contradicciones de su teoría de la aritmética, creyó que la aritmética se tambaleaba, los seguidores de Berkeley y Hume se inclinan a creer que la realidad se tambalea. Y mientras admiran a la ciencia, su filosofía de la ciencia les lleva por necesidad a conclusiones que equivalen a la perspectiva de que la ciencia se tambalea. Para ellos, lo que los científicos consideran como los mayores descubrimientos de la ciencia, los descubrimientos de leyes, no son sino trucos, o reglas de transformación, o si no, nada más que modos de hablar superfluos aunque quizá bastante eficaces.

Esas consecuencias no buscadas deben considerarse como otras tantas refutaciones de las teorías del conocimiento que les dieron origen.

Toda crítica contra teorías rivales debería usarse siempre con la intención de encontrar los puntos débiles de la propia posición. Pero yo no encuentro ninguna consecuencia no buscada ni resultados absurdos que surjan de mi enfoque. Según mi enfoque, es razonable aceptar las concepciones del sentido común siempre que resistan las críticas: la ciencia surge de la crítica y del sentido común y de la imaginación. (Por ejemplo, creo, como el sentido común, en la realidad de las cosas materiales y, por tanto, de la materia. No soy «materialista», sin embargo; no sólo porque creo en las mentes, sino porque no creo que la doctrina de que la materia es última e inexplicable haya resistido a las críticas[113]. Además, es una doctrina insulsa

[113] [Véase la exposición de Popper sobre la historia de la idea de materia en su «Epílogo metafísico», tomo III del *Post Scriptum:* en *El yo y su cerebro (S. & I. B.)*, capítulos 1 y 3; y en «Philosophy and Physics», *Atti del XII Con-*

y sin imaginación. Si un día somos capaces de explicar la materia de forma satisfactoria —por ejemplo, como una perturbación en un campo de fuerzas o de propensiones— debemos cuidarnos mucho de pensar que nos hemos *deshecho* de ella, como muchos «espiritualistas» se inclinan a creer aún hoy. Al explicar un eclipse o una tormenta o una bomba atómica no mostramos que son irreales o meras apariencias. Al afirmar que un coche o un trozo de queso es una estructura formada por partes que no son coches o trozos de queso, no queremos decir que los coches o los trozos de queso sean ilusiones. Sin embargo, por extraño que parezca, hasta Russell habla como si la física, al explicar la estructura de la materia, hubiera mostrado que la materia era una ilusión)[114].

Admitamos que hay personas que consideran perversas algunas de mis teorías, pero estas conclusiones mías «perversas» tienen todas el carácter de rechazos de creencias *filosóficas* ampliamente difundidas. Así, se ha considerado una perversidad de mi parte el resaltar la importancia de la falsación en vez de la importancia de la verificación. Esto, dijo Schlick, es perverso porque queremos estar en lo cierto, no equivocados. No veo razón alguna para disputar el argumento psicológico de Schlick. Yo también prefiero estar en lo cierto; y precisamente por eso es por lo que prefiero corregirme a mí mismo o, si es necesario, que me corrijan otros. Me esfuerzo, pues, en detectar las faltas de mis propios argumentos, es decir, criticarlos, refutarlos. En general, nuestras contrastaciones son intentos de falsación precisamente porque queremos acercarnos a la verdad y evitar la falsedad: porque preferimos estar en lo cierto. Otros se han quejado también, como Schlick, de que no tomo en cuenta los argumentos positivos o en apoyo (en contraposición a las refutaciones). Pero esta queja revela un prejuicio inductivista. La idea principal de la inducción y la verificación es buscar ejemplos sustentadores. Con esta idea filosófica o metodológica estoy en total desacuerdo: los ejemplos meramente sustentadores son por lo general demasiado fáciles para que merezca la pena tenerlos; se pueden tener con sólo pedirlos; por tanto, no

---

gresso Internazionale di Filosofia, Venecia, 1958 (Florencia, 1960) 2, pp. 367-74. *Ed.*]

[114] Véase, por ejemplo, *Portraits from Memory*, 1956, p. 145: «La verdad es, desde luego, que la mente y la materia son, ambas, ilusiones. Los físicos descubren este hecho sobre la materia...». Yo me inclino a afirmar lo contrario. Nunca antes habíamos tenido argumentos científicos (frente a los de sentido común) tan fuertes en favor de la realidad de la materia como los que tenemos ahora, cuando hemos empezado a entenderla. Se ha transformado, de una dudosa idea metafísica sobre la que no podíamos saber nada, en una estructura sobre la que pueden hacerse muchas preguntas claras y enormemente interesantes e, incluso, quizá puedan contestarse.

son de ningún peso; y cualquier apoyo capaz de tener peso sólo puede descansar en contrastaciones ingeniosas, emprendidas con la intención de refutar nuestra hipótesis, si puede ser refutada.

También se ha sugerido que mi intento de rechazar la concepción popular de que «la ciencia comienza por la observación» es perverso, si no se trata de un intento insincero de impresionar a la gente diciendo cosas nuevas y paradójicas. Pero esa concepción es solamente parte de una *filosofía* muy popular, de una teoría inductivista del conocimiento; y su negación es mucho menos perversa que el choque con el sentido común al que dan lugar las consecuencias de esta teoría tan popular. Además, los principales dogmas del inductivismo popular tienen que abandonarse cuando uno se da cuenta de que quizá ni siquiera la teoría de Newton sea verdadera, a pesar de lo increíblemente fuerte que es la evidencia empírica en su apoyo.

No puedo encontrar ninguna consecuencia de mi propia teoría que choque con el sentido común ni con la ciencia; tampoco mis críticos han encontrado esos choques, aunque han encontrado gran cantidad de choques con las doctrinas filosóficas aceptadas. Pero la teoría subjetivista del conocimiento, y con ella la interpretación instrumentalista de las teorías científicas, choca no sólo con el sentido común, sino también con la ciencia y con la tradición racionalista. Rechaza a éstos, aunque sin pretenderlo. Y del mismo modo que los rechaza, puede ser rechazada por ellos.

Con esto concluyo mi crítica de los argumentos en favor del idealismo. Pero todavía tengo que decir algo sobre el realismo, tanto en su favor, como a modo de crítica.

15. EL OBJETIVO DE LA CIENCIA

> *No podría asignarse destino mejor a una teoría física que el de señalar, ella misma, el modo de introducir una teoría más comprensiva dentro de la cual pueda seguir vigente como caso restrictivo.*
>
> ALBERT EINSTEIN

Hasta ahora he argüido en apoyo de la ciencia sobre todo a través de la crítica del idealismo. Voy a ofrecer ahora algunos argumentos positivos en favor del realismo antes de proceder, en la próxima sección, a señalar algunas de sus dificultades.

Los argumentos positivos que tengo en mente se apoyan en la relación entre el realismo y el objetivo de la ciencia.

Desde la publicación de *Logik der Forschung* (es decir, desde 1934) he desarrollado un tratamiento más sistemático del problema del método científico: he tratado de comenzar con algunas sugerencias sobre los objetivos de la actividad científica y de derivar la mayor parte de lo que tengo que decir sobre los métodos de la ciencia —incluidos muchos comentarios sobre su historia— de esta sugerencia. Aquí me limitaré a explicar la sugerencia y a señalar su conexión con los problemas del realismo.

Hablar de «el objetivo» de la actividad científica quizá pueda sonar algo ingenuo; porque está claro que los diferentes científicos tienen diferentes objetivos y que la ciencia misma (sea lo que fuere lo que eso significa) no tiene objetivos. Todo esto lo admito. Pero cuando hablamos de ciencia parecemos sentir, más o menos claramente, que hay algo característico de la actividad científica; y puesto que la actividad científica se parece mucho a una actividad racional, y puesto que una actividad racional debe tener un objetivo, el intento de describir el objetivo de la ciencia puede no ser enteramente fútil.

Sugiero que el objetivo de la ciencia es encontrar *explicaciones satisfactorias* de cualquier cosa que nos parezca a nosotros que necesita explicación. Por una *explicación* (o una explicación causal) se entiende un conjunto de enunciados uno de los cuales describe el estado de las cosas a explicar (el *explicandum*) mientras que los otros, los enunciados explicativos, forman la «explicación», en el sentido más estricto de la palabra (el *explicans* del *explicandum*).

Podemos considerar, por lo general, que se sabe más o menos que el *explicandum* es verdadero o se supone que se sabe. Porque tiene poco sentido pedir una explicación de un estado de cosas que puede resultar ser totalmente imaginario. (Los platillos volantes pueden representar ese caso: la explicación que se necesita puede no ser sobre platillos volantes, sino sobre los relatos sobre platillos volantes; pero si resultase que existen los platillos volantes, entonces ya no sería necesaria la explicación de los *relatos.*) Por otro lado, el *explicans,* que es el objeto de nuestra investigación, no será conocido, por regla general: tendrá que ser descubierto. Así la explicación científica, siempre que sea un descubrimiento, será *la explicación de lo conocido por lo desconocido*[115].

[115] Véase el último párrafo de mi texto (antes de la cita final) de mi «Nota sobre Berkeley como precursor de Mach», reimpresa en *El desarrollo del conocimiento científico (C. & R.).* capítulo 6.

El *explicans*, para ser satisfactorio (el carácter de satisfactorio puede ser cuestión de grados), debe satisfacer una serie de condiciones. En primer lugar, debe entrañar lógicamente al *explicandum*. En segundo lugar el *explicans* tiene que ser verdadero, aunque, en general, no se sabrá que es verdadero; en cualquier caso, no debe saberse que es falso, ni siquiera después del examen más crítico. Si no se sabe que es verdadero (como ocurrirá normalmente), debe haber evidencia *independiente* en su favor. Dicho de otro modo, debe ser contrastable *independientemente*; y será tanto más satisfactorio cuanto mayor sea la severidad de las contrastaciones independientes a las que ha sobrevivido.

Todavía tengo que aclarar mi uso de la expresión «independiente», con sus opuestos, «*ad hoc*» y (en casos extremos) «circular».

Sea *a* un *explicandum* del que se sabe que es verdadero. Puesto que *a* se sigue trivialmente de *a* misma, siempre que podemos ofrecer *a* como explicación de sí misma. Pero esto sería sumamente insatisfactorio, aunque en este caso sabríamos que el *explicans* es verdadero y que el *explicandum* se sigue de él. *Por tanto, tenemos que excluir explicaciones de este tipo a causa de su circularidad.*

Sin embargo, el tipo de circularidad que tengo en mente aquí es cuestión de grado. Consideremos el siguiente diálogo: «¿Por qué está el mar tan agitado hoy?» —«Porque Neptuno está muy enfadado». —«¿Con qué evidencia apoyas tu enunciado de que Neptumo está muy enfadado?» —«Bueno, ¿no *ves* lo *muy* agitado que está el mar? ¿Y no está siempre agitado cuando se enfada Neptuno?». Esta explicación se encuentra insatisfactoria porque (como en el caso de la explicación totalmente circular) la única evidencia en apoyo del *explicans* es el propio *explicandum*[116]. La sensación de que este tipo de explicación casi circular o *ad hoc* es sumamente insatisfactoria y el correspondiente requisito de que las explicaciones de este tipo se eviten están, creo, entre las principales fuerzas en el desarrollo de la ciencia: la insatisfacción está entre los primeros frutos del enfoque crítico o racional.

Para que el *explicans* no sea *ad hoc* tiene que ser rico en contenido: debe tener una variedad de consecuencias contrastables que sean diferentes del *explicandum*. Son esas diferentes consecuencias contrastables en las que estoy pensando cuando hablo de contrastaciones *independientes* o de evidencia *independiente*.

---

[116] Este razonamiento ha sobrevivido en Thales (Diels-Kranz 10, vol. I, p. 456, línea 35); Anaximandro (D.-K. A11, A28); Anaxímenes (D.-K. A17, B1).

Aunque estas observaciones quizá puedan ayudar a aclarar algo la idea intuitiva de un *explicans* independiente contrastable, todavía son insuficientes para caracterizar una explicación satisfactoria e independientemente contrastable. Por que si *a* es nuestro *explicandum* —sea *a* de nuevo «el mar está agitado hoy»— entonces siempre podemos ofrecer un *explicans* altamente insatisfactorio que sea completamente *ad hoc* aunque tenga consecuencias independientemente contrastables. Todavía podemos escoger esas consecuencias a voluntad. Porque podemos escoger por ejemplo «estas ciruelas son jugosas» y «todos los cuervos son negros». Sea *b* su conjunción. Entonces podemos tomar como *explicans* simplemente la conjunción de *a* y *b*: satisfará todos los requisitos que hemos enunciado hasta ahora, pero será *ad hoc* y totalmente insatisfactorio intuitivamente.

Sólo si exigimos que las explicaciones utilicen leyes universales de la naturaleza (complementadas con las condiciones iniciales) podemos progresar hacia la realización de la idea de explicaciones independientes, no *ad hoc*. Porque las leyes universales de la naturaleza *pueden* ser enunciados ricos en contenido, de forma que *pueden contrastarse independientemente* en todo lugar y en todo momento. Por tanto, si se utilizan como explicaciones, no *pueden* ser *ad hoc* porque pueden permitirnos interpretar el *explicandum* como un ejemplo de un efecto reproducible. Todo esto es verdad, sin embargo, sólo si nos limitamos a leyes universales que sean contrastables, es decir, falsables. Es aquí donde entran en juego el problema de la demarcación y el criterio de falsabilidad.

La pregunta «¿qué clase de explicación puede ser *satisfactoria*?» lleva, pues, a la respuesta: una explicación en términos de leyes universales y condiciones iniciales contrastables y falsables. Una explicación de este tipo será tanto más satisfactoria cuanto más contrastables sean esas leyes y cuanto mejor contrastadas hayan sido. (Esto es también lo aplicable a las condiciones iniciales.)

De este modo, la conjetura de que es objetivo de la ciencia encontrar explicaciones satisfactorias nos lleva más lejos, a la idea de perfeccionar el grado de satisfacción que procuran nuestras explicaciones mejorando su grado de contrastabilidad, es decir, avanzando hacia explicaciones que se contrasten mejor; lo que significa —como se muestra en los capítulos VI y VIII de *L. I. C.*— avanzar hacia teorías de un contenido cada vez más rico, de un grado superior de universalidad, de un grado superior de precisión. Esto, sin duda, está plenamente de acuerdo con la historia y la práctica real de las ciencias teóricas.

Podemos llegar a un resultado fundamentalmente similar de

otro modo. Si es el objetivo de la ciencia explicar, entonces será también su objetivo explicar lo que hasta el momento se ha aceptado como un *explicans*, tal como una ley de la naturaleza. Así, la tarea de la ciencia se renueva constantemente. Podríamos seguir eternamente, avanzando hacia explicaciones de un nivel cada vez mayor de universalidad, a menos, naturalmente, que llegásemos a una *explicación última*: es decir, a una explicación que no tuviese ni necesitase explicación ulterior.

¿Existen explicaciones últimas? La doctrina que he denomidado «esencialismo» sostiene la perspectiva de que la ciencia debe buscar explicaciones últimas en términos de esencias[117]: si podemos explicar el comportamiento de una cosa en términos de su esencia —de sus propiedades esenciales— entonces no puede plantearse ninguna cuestión ulterior (excepto, quizá, la cuestión teológica del Creador de las esencias). Así, Descartes creía que había explicado la física en términos de *la esencia de un cuerpo físico* que, enseñaba él, era extensión; y algunos newtonianos, siguiendo a Roger Côtes, creían que *la esencia de la materia* era su inercia y su capacidad de atraer a otra materia y que la teoría de Newton podía derivarse de esas propiedades esenciales de toda la materia y ser explicada así, en último término, por ellas. El propio Newton era de diferente opinión. Era una hipótesis concerniente a la explicación causal última o esencial de la gravedad misma, lo que él tenía en mente cuando escribió en el *Scholium generale* al final de los *Principia:* «Hasta ahora he explicado los fenómenos... por la fuerza de la gravedad, pero aún no he averiguado *la causa de la gravedad misma*... y yo no invento hipótesis arbitrariamente [es decir, *ad hoc*]»[118].

Yo no creo en la doctrina de la explicación última. En el pasado, por regla general, los críticos de esta doctrina han sido instrumentalistas: interpretaban las teorías científicas como *nada más que* instrumentos de predicción sin ninguna capacidad explicativa. Tampoco estoy de acuerdo con ellos. Pero hay una tercera posibilidad. Una «tercera concepción» como la he lla-

---

[117] He discutido y criticado el esencialismo más extensamente en mi artículo «Tres concepciones sobre el conocimiento humano», *El desarrollo del conocimiento científico (C. & R.)*, capítulo 3, donde también hago referencia a exposiciones anteriores (en la última nota de la sección II). [El esencialismo y la exigencia de explicaciones últimas están también naturalmente, íntimamente relacionadas con el justificacionismo. Véase sección 2 más arriba. *Ed.*]

[118] Véanse también las cartas de Newton a Richard Bentley, del 17 de enero y, sobre todo, del 25 de febrero de 1693 («1692-3»). He citado pasajes de esta carta en la sección III de «Tres concepciones sobre el conocimiento humano», donde se expone un poco más extensamente el problema.

mado. Ha sido bien descrita como un «esencialismo modificado», con el énfasis de la palabra «modificado»[119].

Esta «tercera concepción» que yo sostengo, modifica el esencialismo de manera radical. En primer lugar, yo rechazo la idea de explicación última. Mantengo que toda explicación puede ser más explicada por una teoría de mayor universalidad. No puede haber explicación que no necesite otra explicación más, porque ninguna puede ser una descripción autoexplicativa de una esencia (tal como una definición esencialista de cuerpo, como sugirió Descartes). En segundo lugar, rechazo las *preguntas de «¿qué es?»*: preguntas sobre qué es una cosa, cuál es su esencia, o su verdadera naturaleza. Porque debemos renunciar a la concepción, característica del esencialismo, de que en cada cosa singular hay una esencia, una naturaleza o principio inherente (tal como el espíritu del vino en el vino) que hace necesariamente que sea lo que es y, por tanto, que actúe como lo hace. Esta concepción animista no explica nada; pero ha llevado a los esencialistas (como Newton) a rehuir las propiedades relacionales, tales como la gravedad y creer, con argumentos considerados válidos *a priori*, que una explicación satisfactoria tiene que ser en términos de propiedades inherentes (como opuestas a las propiedades relacionales). La tercera y última modificación del esencialismo es ésta: debemos renunciar a la concepción, estrechamente relacionada con el animismo (y característica de Aristóteles, a diferencia de Platón) de que es a las propiedades esenciales inherentes a *cada individuo o cosa singular* a las que hay que recurrir como explicativas del comportamiento de esa cosa. Porque esta concepción no arroja ninguna luz, en absoluto, sobre por qué cosas individuales diferentes se comportan del mismo modo. Si se dice que «porque sus esencias son iguales» surge la pregunta «*¿por qué no podrían existir tantas esencias distintas como cosas distintas hay?*».

Platón trató de resolver precisamente este problema diciendo que las cosas individuales iguales son retoños, y, por tanto, copias, de la misma «Forma» original, que es, por tanto, algo «exterior» y «anterior» y «superior» a las diversas cosas individuales; y, efectivamente, todavía no tenemos una teoría

[119] El término «esencialismo modificado» se usó como descripción de mi «tercera concepción» por un reseñista de «Tres concepciones sobre el conocimiento humano», en *The Times Literary Supplement* 55, 1956, p. 527. Para evitar malentendidos, quiero decir aquí que mi aceptación del término no debe interpretarse como una concesión a las doctrinas de «realidad última» y «explicación última» y menos aún como una concesión a la doctrina de las definiciones esencialistas. Me adhiero a las críticas de esta doctrina que he expuesto en *La sociedad abierta (O. S.)*, cap. 11, sección II (sobre todo, nota 42), y en otros lugares, véase más abajo, sección 31; también nota 2 a la sección 19 de *L. I. C.*

mejor de la similaridad. Aun hoy apelamos a su origen común si queremos explicar el parecido de dos hombres, o de un pájaro y un pez o de dos camas o dos coches o dos lenguas o dos procedimientos legales; es decir, explicamos la similaridad, en general, genéticamente, y si de esto hacemos un sistema metafísico, es probable que se convierta en una filosofía historicista. La solución de Platón fue rechazada por Aristóteles; pero como la solución de Aristóteles del esencialismo no contiene ni un indicio de solución, parece que nunca captó el problema[120].

Al escoger explicaciones en términos de leyes universales de la naturaleza, ofrecemos una solución precisamente a este último problema (platónico). Porque concebimos a todas las cosas individuales y a todos los hechos individuales como sujetos a esas leyes. Las leyes (que, a su vez, *necesitan* explicación ulterior) explican, pues, las regularidades o similaridades entre cosas individuales o hechos o sucesos singulares. Y esas leyes no son inherentes a las cosas singulares (ni son tampoco ideas platónicas fuera del mundo). Las leyes de la naturaleza se conciben, por el contrario, como descripciones (conjeturables) de las propiedades estructurales ocultas en la naturaleza, de nuestro propio mundo.

Aquí está entonces la similaridad entre mi propia concepción (la «tercera concepción») y el esencialismo: aunque no pienso que vayamos a describir nunca, por medio de nuestras leyes universales, una esencia *última* del mundo, no dudo que podamos intentar sondear, cada vez más profundamente, la estructura de nuestro mundo o, como podríamos decir, propiedades del mundo que serán cada vez más esenciales, o de mayor profundidad.

Cada vez que procedemos a explicar alguna ley o teoría conjetural de un grado superior de universalidad, estamos descubriendo más cosas sobre el mundo: estamos penetrando más profundamente en sus secretos. Y cada vez que conseguimos falsar una teoría de este tipo hacemos un nuevo descubrimiento importante. Porque esas falsaciones son de la mayor importancia. Nos enseñan lo inesperado. Y nos tranquilizan respecto a nuestras teorías que, aunque son hechas por nosotros, aunque son invenciones nuestras, no por ello dejan de ser auténticas aserciones sobre el mundo; porque pueden *chocar* con algo que nosotros no hemos hecho.

[120] En cuanto a la teoría de Platón sobre las formas o ideas, «una de sus funciones más importantes [es] explicar la similaridad de las cosas sensibles...». *Cf. La sociedad abierta (O. S.)*, cap. 3, sección V; véanse también notas 19 y 20 y texto. Ahí se menciona la incapacidad de la teoría de Aristóteles para cumplir esa función (en la 3.ª edición, 1957) al final de la nota 54 al capítulo 11.

Nuestro «esencialismo modificado» es, creo, de gran ayuda cuando se plantea la cuestión de la forma lógica de las leyes naturales. Sugiere que nuestras leyes o teorías deben ser *universales*, es decir, deben formular aserciones sobre todas las regiones espacio-temporales del mundo. Sugiere, además, que nuestras teorías formulan aserciones sobre propiedades estructurales relacionales del mundo; que las propiedades descritas por una teoría explicativa deben ser, en uno u otro sentido, más profundas que las explicadas.

Estas dos ideas —la de las propiedades estructurales o relacionales del mundo y la de la profundidad de una teoría— necesitan aclaración.

A menudo explicamos el comportamiento legiforme de ciertas cosas individuales en términos de su *estructura*. Podemos explicar así, y entender, el funcionamiento de un reloj después de desmontarlo unas cuantas veces y de volverlo a montar; porque de ese modo podemos aprender a entender su estructura y su forma de funcionar como consecuencia de su estructura. Ahora bien, si consideramos un poco más atentamente este procedimiento, encontraremos que, en una explicación estructural de este tipo, siempre presuponemos algún comportamiento legiforme diferente del que vamos a explicar (y «más profundo» que él). Por ejemplo, lo que queremos explicar, en el caso del reloj, es el movimiento regular de sus ruedas y sus agujas. Lo hacemos analizando su estructura; pero también tenemos que suponer que las distintas piezas que componen la estructura son *rígidas* (es decir, que conservan sus formas geométricas y sus «extensiones») y que son *impenetrables* (es decir, que se empujan unas a otras —si una pieza se pone en el camino de otra— en lugar de moverse una pieza a través de la otra, por así decirlo). Esas dos propiedades legiformes, la rigidez y la impenetrabilidad de ciertos cuerpos, pueden, a su vez, ser explicadas de nuevo estructuralmente, por ejemplo, por *retículas de átomos* que, se ha conjeturado, constituyen la estructura material de este tipo de cuerpo. Pero en esta segunda explicación no sólo conjeturamos que ciertas partes —los átomos— están dispuestas en una estructura reticular, sino que además suponemos que se cumplen ciertas leyes de atracción y repulsión entre los átomos. Esas, a su vez, pueden explicarse por la estructura subatómica de los átomos junto con las leyes que gobiernan el comportamiento de las partículas subatómicas y así sucesivamente. Todo esto puede expresarse con la metáfora, de admitida vaguedad, de que las leyes de la naturaleza enuncian «propiedades estructurales del mundo». (La metáfora es vaga precisamente porque, a cualquier nivel, no explica sólo la estructura, sino también las leyes; pero es permisible porque,

a cualquier nivel, las leyes se explican en parte por las estructuras y también porque es por lo menos concebible que, a algún nivel, estructura y ley puedan llegar a ser indistinguibles, que las leyes *impongan* un cierto tipo de estructura al mundo y que puedan interpretarse, alternativamente, como *descripciones* de esta estructura[121]. Esto es lo que parecen proponerse, si no lo han conseguido ya, las teorías de campos de la materia.) Y nada más sobre la idea de estructura.

La segunda idea que necesita clarificación es la de «*profundidad*». Desafía, creo yo, cualquier tentativa de análisis lógico exhaustivo; es, sin embargo, una guía para nuestras intuiciones. (Esto ocurre en matemáticas; en presencia de los axiomas, todos sus teoremas son lógicamente equivalentes y sin embargo hay una gran diferencia en la «profundidad», que difícilmente es susceptible de análisis lógico)[122]. La «profundidad» de una teoría científica parece estar muy estrechamente relacionada con su simplicidad y, por tanto, con la riqueza de su contenido. (No ocurre lo mismo con la profundidad de un teorema matemático, cuyo contenido puede considerarse nulo.) Parecen requerirse dos ingredientes: un contenido rico y una cierta coherencia o compacidad (u «organicidad») del estado de cosas que se describe. Es este último ingrediente el que es muy difícil de analizar, aunque es intuitivamente bastante claro, y es lo que los esencialistas trataban de describir cuando hablaban de esencias, en contraposición a meras acumulaciones de propiedades accidentales. No creo que podamos hacer mucho más que referirnos aquí a una idea intuitiva, ni que necesitemos hacer mucho más. Porque en el caso de cualquier teoría deter-

[121] En el Apéndice *10 de *L. I. C.*, trato de ciertos problemas relativos a las leyes de la naturaleza y a su carácter dual, debido al cual son al mismo tiempo, en un sentido, necesarias (y a ello alude aquí la palabra «imponen») y contingentes (a lo que alude la palabra «descripción»). Para el problema de explicar la rigidez («extensión») y la impenetrabilidad, véase también el «Epílogo metafísico» en el tomo III del *Post Scriptum*, secciones 21 y 27, en que pueden encontrarse algunos comentarios sobre la teoría de campos de la materia.

[122] Se ha sugerido que, en matemáticas, la profundidad de un teorema puede medirse suponiendo que *a)* la profundidad aumenta con el número de pasos —es decir, con la longitud— de la prueba más corta, y que *b)* disminuye con la longitud del propio teorema. Esto haría relativa la idea de profundidad, ya que un cierto teorema t, que en el lenguaje formalizado L_1 es corto y que necesita una prueba muy larga en L_1, puede ser una fórmula muy larga en el lenguaje L_2 y puede, sin embargo, servir como un axioma de L_3. Pero intuitivamente parece haber algo absoluto respecto a la profundidad, una cualidad que claramente se pierde con la medida propuesta, y que tiene algo que ver con *a)* la profundidad de la idea o el carácter de la prueba (más sencilla dentro del sistema más sencillo) más que con su longitud y *b)* la fertilidad o aplicabilidad general de esta idea, como método de prueba, más que con la brevedad del teorema.

minada que se proponga, es la riqueza de su contenido y, por tanto, su grado de contrastabilidad, lo que decide su interés, y el resultado de las contrastaciones efectivas, lo que decide su destino. Desde el punto de vista del método, podemos mirar a su profundidad, su coherencia, incluso a su belleza, como simples guías o estímulos para nuestra intuición y nuestra imaginación.

No obstante, parece haber algo como una *condición suficiente* para la profundidad o para los grados de profundidad, que puede analizarse lógicamente. Trataré de explicar eso con ayuda de un ejemplo sacado de la historia de la ciencia.

Es bien sabido que la dinámica de Newton logró una unificación de la física terrestre de Galileo y de la física celeste de Kepler. Se dice con frecuencia que la dinámica de Newton puede inducirse de las leyes de Galileo y de Kepler y se ha afirmado incluso que puede deducirse estrictamente de ellas[123]. Pero no es así: desde un punto de vista lógico, la teoría de Newton, estrictamente hablando, contradice tanto la de Galileo como la de Kepler (aunque estas teorías pueden obtenerse, desde luego, como aproximaciones, una vez que tenemos la teoría de Newton para operar). Por esta razón es imposible derivar la teoría de Newton de la de Galileo ni de la de Kepler, ni de ambas, sea por inducción o por deducción. Porque nunca puede proceder una inferencia, ni deductiva ni inductiva, de premisas consistentes a una conclusión que contradice formalmente esas premisas.

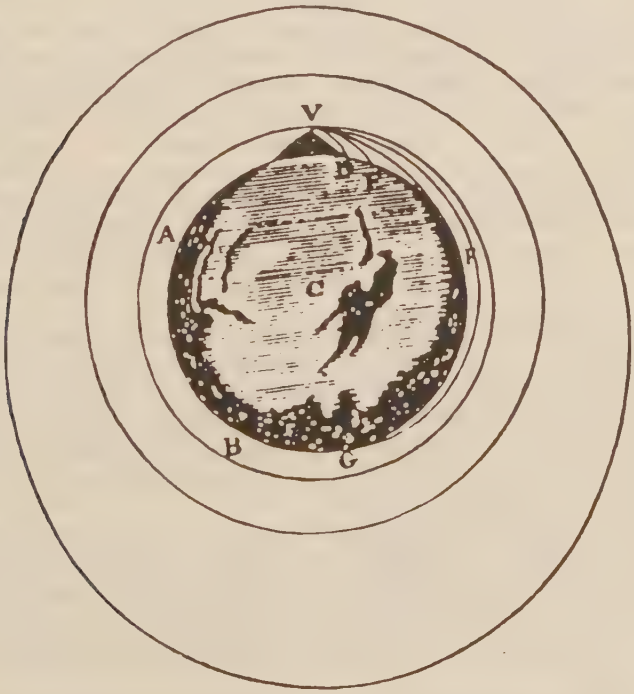
Considero que éste es un argumento muy fuerte contra el inductivismo.

Indicaré ahora brevemente las contradicciones entre la teo-

---

[123] Lo que puede deducirse de las leyes de Kepler (véase Max Born, *Natural Philosophy of Cause and Chance*, 1949, pp. 129-133) es que para todos los planetas, la aceleración hacia el Sol es igual, en cualquier momento, a k/r^2, donde r es la distancia en ese momento entre el planeta y el Sol y k una constante, la misma para todos los planetas. Sin embargo, este resultado mismo contradice formalmente la teoría de Newton (excepto si se supone que las masas de los planetas son todas iguales o, si son desiguales, infinitamente pequeñas en comparación con la masa del Sol). Este hecho es consecuencia de lo que se dice aquí, en el texto que sigue a la nota 125, sobre la tercera ley de Kepler. Pero, además, hay que recordar que ni la teoría de Kepler ni la de Galileo contienen el concepto de Newton de *fuerza*, que tradicionalmente se introduce en estas deducciones sin más discusión; como si este concepto («oculto») se desprendiera de los hechos en lugar de ser el resultado de una nueva interpretación de los hechos (es decir, de los «fenómenos» descritos por las leyes de Kepler y Galileo) a la luz de una teoría totalmente nueva. Sólo *después* de introducir el concepto de *fuerza* (e incluso la proporcionalidad de la masa gravitatoria e inerte) es posible relacionar la fórmula de la aceleración con la ley inversa del cuadrado, la ley de la atracción de Newton (por medio de un supuesto como el de que las masas de los planetas son despreciables).

ría de Newton y las de sus predecesores. Galileo afirma que una piedra lanzada, o un proyectil, se mueve describiendo una parábola, excepto en el caso de una caída libre vertical, en que se mueve con aceleración constante en línea recta. (Despreciaremos la resistencia del aire en toda esta discusión.) Desde el punto de vista de la teoría de Newton, esas dos aserciones son falsas, por dos razones distintas. La primera es falsa porque la trayectoria de un proyectil de largo alcance, tal como un cohete intercontinental (lanzado en dirección horizontal o ascendente) no será parabólica ni siquiera aproximadamente: será elíptica. Se convierte aproximadamente en una parábola sólo si la distancia total del vuelo del proyectil es despreciable en comparación con el radio de la Tierra. Este argumento lo utilizó el propio Newton en los *Principia*, así como en su versión popular, *The System of the World*, donde lo ilustra con ayuda de la figura siguiente[124].



[124] Véase Newton, *Principia*, el Escolio al final de la sección II del Libro I; p. 55, de la edición de 1934 (traducción [inglesa] de Motte, revisada por Cajori). La figura, sacada de *The System of the World*, y la cita que se da aquí se encuentran en la p. 551 de dicha edición.

La figura de Newton ilustra su enunciado de que si la velocidad del proyectil se aumenta, y con ella la distancia de su vuelo, «al fin, excediendo los límites de la tierra... [pasará] al espacio sin tocarla». Y describirá, aproximadamente, una elipse kepleriana durante todo su trayecto.

Por tanto, un proyectil se mueve sobre la Tierra describiendo una elipse con excentricidad finita, y no una parábola. Naturalmente, para lanzamientos suficientemente cortos, una parábola será una aproximación excelente; pero el recorrido parabólico no es deducible extrictamente de la teoría de Newton a menos que añadamos a ésta una condición inicial *falsa* de hecho (y que, por cierto, es irrealizable en la teoría de Newton, puesto que lleva a consecuencias totalmente absurdas) en el sentido de que el radio de la Tierra sea infinito. Si no admitimos ese supuesto, que *sabemos que es falso*, entonces siempre obtendremos una elipse, una contradicción de la ley de Galileo, según la cual deberíamos obtener una parábola.

Una situación lógica exactamente análoga surge en relación con la segunda parte de la ley de Galileo que afirma la existencia de una aceleración *constante*. Desde el punto de vista de la teoría de Newton, la aceleración de los cuerpos en caída libre nunca es constante: siempre aumenta durante la caída, debido al hecho de que el cuerpo va aproximándose cada vez más al centro de atracción. Este efecto es muy considerable si el cuerpo cae desde gran altura pero despreciable, naturalmente, si la altura es despreciable en comparación con el radio de la Tierra. En este caso, podemos obtener la teoría de Galileo de la de Newton si introducimos de nuevo la *falsa* suposición de que el radio de la Tierra es infinito (o la altura de la caída, cero).

Las contradicciones que he señalado están lejos de ser despreciables para los cohetes de largo alcance. A éstos podemos aplicarles la teoría de Newton (con correcciones para la resistencia del aire, naturalmente) pero no la de Galileo: ésta lleva simplemente a resultados falsos, como puede mostrarse con ayuda de la teoría de Newton.

Con respecto a las leyes de Kepler, la situación es similar. Es obvio que, en la teoría de Newton, las leyes de Kepler son válidas sólo aproximadamente —es decir, estrictamente inválidas— sitenemos en cuenta la atracción mutua entre los planetas[125]. Pero hay contradicciones entre las dos teorías que son

[125] Véase, por ejemplo, P. Duhem, *The Aim and Structure of Physical Theory*, 1905, versión inglesa de P. P. Wiener, 1945, Parte II, cap. IV, sección 4. Duhem dice de modo más explícito lo que está implícito en el enunciado del propio Newton (*Principia*, libro I, proposición LXV, teorema XXV); porque Newton expresa con bastante claridad que siempre que interactúan más de dos cuerpos, las dos primeras leyes de Kepler serán, en el mejor de los casos, sólo

más fundamentales que ésta, algo obvia. Porque, incluso aunque despreciemos, como concesión a nuestros oponentes, la atracción mutua entre los planetas, la tercera ley de Kepler, considerada desde el punto de vista de la dinámica de Newton, no puede ser más que una aproximación que es aplicable a un caso muy especial: a planetas cuyas masas sean todas iguales o, si son desiguales, despreciables en comparación con la masa del Sol. Como esto no ocurre ni siquiera aproximadamente con dos planetas, si uno de ellos es muy ligero y el otro muy pesado, está claro que la tercera ley de Kepler contradice la teoría de Newton de la misma forma precisamente en que lo hace la de Galileo.

Esto puede mostrarse fácilmente como sigue: La teoría de Newton da, para un sistema de dos cuerpos —un sistema estelar binario—, una ley que los astrónomos llaman a menudo «la ley de Kepler», puesto que se relaciona estrechamente con la tercera ley de Kepler. Esta llamada «ley de Kepler» dice que si m_0 es la masa de uno de los dos cuerpos —del Sol, por ejemplo— y si m_1 es la masa del otro cuerpo —un planeta, por ejemplo— entonces, escogiendo unidades de medidas apropiadas, podemos derivar de la teoría de Newton

$$1) \qquad a^3/T^2 = m_0 + m_1$$

donde a es la distancia media entre los dos cuerpos y T el tiempo de una revolución completa. Ahora la tercera ley del propio Kepler afirma que

$$2) \qquad a^3/T^2 = \textit{constante}$$

es decir, la misma constante para *todos* los planetas del sistema solar. Está claro que obtenemos esta ley de 1) sólo en el supuesto de que $m_0 + m_1 = \textit{constante;}$ y puesto que $m_0 = \textit{constante}$ para nuestro sistema solar si identificamos a m_0 con la masa del Sol, obtenemos 2) de 1) siempre que supongamos que m_1 es la misma para todos los planetas; o si esto es *falso* de hecho (como lo es, en efecto, ya que Júpiter es varios miles de veces mayor que los planetas más pequeños), que las masas de los planetas *son todas cero en comparación con la del Sol,* de modo que podamos decir que $m_1 = 0$, *para todos los planetas.* Esta es una aproximación bastante buena desde el punto de vista de la teoría de Newton; pero al mismo tiempo, el decir que $m_1 = 0$ no es solo falso, estrictamente hablando, sino irrea-

válidas aproximadamente, y esto, incluso, sólo en casos muy especiales, de los cuales analiza dos con cierto detalle. Por cierto, la fórmula 1) más abajo, se sigue inmediatamente del libro I, proposición LIX, en vista de la proposición XV del libro I. (Véase también el libro III, proposición XV.) Así, mi propio análisis, como el de Duhem, esta implícito en el de Newton.

lizable desde el punto de vista de la teoría de Newton. (Un cuerpo de masa cero no obedecería las leyes del movimiento de Newton.) Así, incluso aunque olvidemos todo lo referente a la atracción mutua entre los planetas, la tercera ley de Kepler 2) contradice formalmente la teoría de Newton que da 1).

Es importante resaltar que de las teorías de Galileo y Kepler no obtenemos el más ligero indicio de cómo podrían adaptarse estas teorías —qué premisas (falsas) habría que adoptar, o qué condiciones habría que estipular— para interpretarlas por medio de otra teoría de validez más general, tal como la de Newton. *Sólo después de estar en posesión de la teoría de Newton podemos descubrir si, y en qué sentido, las teorías anteriores son aproximaciones a ella.* Podemos expresar este hecho brevemente diciendo que, aunque desde el punto de vista de la teoría de Newton, las de Galileo y Kepler son aproximaciones excelentes de ciertos resultados especiales de Newton, no puede decirse que la teoría de Newton sea, desde el punto de vista de las otras dos teorías, una aproximación a los resultados de ellas. Todo esto muestra que la lógica, sea deductiva o inductiva, no puede dar el paso desde esas teorías a la dinámica de Newton[126]. Sólo el ingenio puede darlo. Una vez que se ha dado, puede decirse que los resultados de Galileo y de Kepler corroboran la nueva teoría.

Aquí, sin embargo, no estoy tan interesado en la imposibilidad de la inducción como en *el problema de la profundidad.* Y en relación con este problema, podemos en efecto aprender algo con nuestro ejemplo. La teoría de Newton unifica las de Galileo y Kepler. Pero, lejos de ser una mera conjunción de esas dos teorías —que representan el papel de *explicanda* en relación con la de Newton— *las corrige al explicarlas.* La tarea explicativa original era la deducción de los resultados anteriores. Pero esta tarea no se cumple deduciendo esos resultados anteriores, sino deduciendo algo mejor en su lugar: nuevos resultados que, bajo las condiciones especiales de los resultados anteriores, se acercan mucho numéricamente a esos resultados anteriores y al mismo tiempo, *los corrigen.* Así, puede decirse que el éxito empírico de la vieja teoría corrobora la nueva teoría; y además, las correcciones pueden ser contrastadas, a su vez, y quizá refutadas o si no, corroboradas. Lo que la situación lógica que he esbozado resalta con fuerza es el hecho de que la nueva teoría no puede ser *ad hoc* o circular. Lejos de repetir su *explicandum,* la nueva teoría lo contradice y lo corrige. De este modo, incluso la evidencia del *explicandum*

[126] Los conceptos de fuerza (*cf.* nota 123) y de acción a distancia introducen nuevas dificultades.

mismo se convierte en evidencia independiente en favor de la nueva teoría. (Por cierto, este análisis nos permite *explicar* —siguiendo líneas similares a las de *L. I. C.*, sección 57— *el valor de las teorías métricas* y de la medición; y nos ayuda, por tanto, a evitar el error de aceptar la medición y la precisión como valores últimos e irreductibles.)

Sugiero que, si, en las ciencias empíricas, una nueva teoría de un nivel superior de universalidad explica con éxito una teoría anterior *corrigiéndola*, entonces esto es un signo seguro de que la nueva teoría ha penetrado más hondo que la vieja. El requisito de que una nueva teoría contenga a la antigua aproximadamente, para valores adecuados de los parámetros de la nueva teoría, puede llamarse (siguiendo a Bohr) «*principio de correspondencia*».

El cumplimiento de este requisito es una condición suficiente de profundidad, como ya he dicho. Que no es una condición necesaria puede verse en el hecho de que la teoría de las ondas electromagnéticas de Maxwell no corrigió, en este sentido, la teoría ondulatoria de la luz de Fresnel. Supone un aumento de profundidad, sin duda, pero en un sentido diferente: «La vieja cuestión de la dirección de las vibraciones de la luz polarizada perdió sentido. Las dificultades relativas a las condiciones límite para la separación de dos medios fueron resueltas por los fundamentos mismos de la teoría. No hacían falta ya hipótesis *ad hoc* para eliminar las ondas luminosas longitudinales. La presión de la luz, tan importante en la teoría de la radiación y que sólo recientemente se ha determinado experimentalmente, podía derivarse como una de las consecuencias de la teoría»[127]. Este brillante pasaje, en el que Einstein esboza algunos de los mayores logros de la teoría de Maxwell y la compara con la de Fresnel, puede tomarse como indicación de que hay otras condiciones suficientes de profundidad que mi análisis no cubre.

La tarea de la ciencia, que, como he sugerido, es encontrar explicaciones satisfactorias, difícilmente puede entenderse si no somos realistas. Porque una explicación satisfactoria es una que no es *ad hoc*; y esta idea —la idea de *evidencia independiente*— difícilmente puede entenderse sin la idea de descubrimiento, de

[127] A. Einstein, *Physikalische Zeitschrift*, 10, 1909, pp. 817 s. El abandono de la teoría de una éter material (implícito en el hecho de que Maxwell no pudiera construir un modelo material satisfactorio de él), puede decirse que da profundidad, en el sentido analizado más arriba, a la teoría de Maxwell, en comparación con la de Fresnel; y esto, me parece a mí, está implícito en la cita del artículo de Einstein. Así pues, quizá la teoría de Maxwell en la formulación de Einstein no constituya realmente un ejemplo de otro sentido de «profundidad». Pero en la forma original de Maxwell pienso que lo es.

progreso hacia niveles más profundos de explicación; sin la idea, por tanto, de que hay algo que tenemos que descubrir; y algo que discutir críticamente.

Y sin embargo, me parece a mí que en metodología no tenemos que presuponer el realismo metafísico. Tampoco podemos obtener ninguna ayuda de él, excepto de tipo intuitivo. Porque una vez que se nos ha dicho que el objetivo de la ciencia es explicar, y que la explicación más satisfactoria será la que sea más rigurosamente contrastable y más rigurosamente contrastada, sabemos todo lo que necesitamos saber como metodólogos. Que el objetivo sea realizable no podemos asegurarlo ni con la ayuda del realismo metafísico ni sin ella, ya que éste sólo puede ofrecernos algún estímulo intuitivo, alguna esperanza, pero ninguna seguridad de ninguna clase. Y aunque puede decirse que un tratamiento racional de la metodología depende de un objetivo, supuesto o conjeturado, de la ciencia, ciertamente no depende de la suposición metafísica, y muy probablemente falsa, de que la verdadera estructura del mundo (si existe) puede enunciarse en el lenguaje humano.

Si la imagen del mundo que traza la ciencia moderna se acerca en algo a la verdad —en otras palabras, si tenemos algo parecido a «conocimiento científico»— entonces las condiciones que prevalecen en casi todo el universo hacen casi imposible el descubrimiento de leyes estructurales del tipo de las que buscamos —y, por tanto, la consecución del «conocimiento científico»—. Porque casi todas las regiones del universo están llenas de radiación caótica, y casi todo el resto, de materia en un estado igualmente caótico. A pesar de esto, la ciencia ha tenido un éxito milagroso en continuar avanzando hacia lo que pienso que debe considerarse su objetivo. Pienso que este extraño hecho no puede explicarse sin demostrar demasidadas cosas (*cf.* sección 3, texto de la nota 19). Pero puede estimularnos a perseguir ese objetivo, aunque no podamos recibir más estímulos ni del realismo metafísico, ni de ninguna otra fuente, para creer que podemos alcanzarlo realmente[128].

**Addendum de 1980 a la sección 15*

1) La sección precedente, como ya se ha dicho, se publicó por primera vez en 1957. Contiene, entre otras cosas, una refu-

[128] La mayor parte de esta sección (15) se publicó por primera vez en *Ratio* I, 1957 (tanto en la edición inglesa como en la alemana). El material se reedita aquí con permiso del editor, mi amigo, el difunto Julius Kraft.

* [También se ha reimprimido en *Conocimiento objetivo (Obj. Kn.)* como capítulo I. La idea que se discute en esta sección de que las teorías pueden *corregir* una ley «observacional» o «fenomenal» que se supone que explican

tación de la concepción que sostenían hombres como Isaac Newton y Max Born, según la cual la teoría de Newton puede derivarse de las leyes de Kepler, sea por un argumento inductivo o por uno deductivo.

Cuando escribí esta sección por primera vez no hice mucho hincapié en la refutación del mito histórico de que la teoría de Newton es el resultado de la inducción porque pensé que había destruido la teoría de la inducción veinte años antes; y era lo bastante optimista como para creer que la resistencia que seguían ofreciendo los defensores de la inducción desaparecería pronto. (No obstante, critiqué con bastante detalle la teoría de Carnap, entonces vigente, de la inducción probabilística, y el resultado fue que él acabó por abandonarla; y la última forma en que defendió la inducción era completamente diferente de la famosa teoría que desarrolló en su extenso pero, en mi opinión, totalmente insostenible libro, *Logical Foundations of Probability*.)

Desde entonces, los inductivistas han cobrado ánimos; en parte porque yo no he replicado ya a sus argumentos, que estaban todos claramente refutados en diversas partes de mis escritos anteriores. Ya no les replicaba porque pensaba, y sigo pensando, que la cuestión estaba decidida desde hacía tiempo y que, por tanto, era ya aburrida.

2) No obstante, puede ser útil repetir aquí, muy brevemente, uno de los argumentos más interesantes contra la inducción, un argumento que está implícito en la sección 15.

Por inducción entiendo un argumento que, dadas ciertas premisas (singulares o particulares), lleva a una conclusión universal, a una teoría universal, o bien con certeza lógica o con «probabilidad» (en el sentido que tiene el término en el cálculo de probabilidades).

El argumento en contra de la inducción que quiero reenunciar aquí es muy sencillo:

Muchas teorías, como la de Newton, que se han conside-

(tal como ocurre, por ejemplo, con la tercera ley de Kepler), la he expuesto repetidas veces en mis conferencias desde los días de Nueva Zelanda. Una de esas conferencias estimuló la corrección de una supuesta ley fenomenal (véase el artículo de 1941 al que me refiero en *La miseria del historicismo* [*Poverty*, 1957, 1960, nota en pp. 134 y s.]). Otra de esas conferencias se publicó en Simon Moser, *Gesetz und Wirklichkeit* (1948), 1949 (véase especialmente p. 57 y ss.) y reimpreso en Hans Albert, *Theorie und Realität*, 1964 (véase especialmente p. 100); se ha publicado también como Apéndice a *Conocimiento objetivo (Obj. Kn.)* con el título de «El cubo y el reflector: dos teorías del conocimiento». Esta misma idea mía fue también el «punto de partida» (como dice él mismo en p. 92) de P. K. Feyerabend, «Explanation, Reduction and Empiricism» (en Herbert Feigl y Grover Maxwell, eds., *Minnesota Studies in the Philosophy of Science* 3, 1962), cuya referencia (66) es a esta sección (tal como se publicó en *Ratio*). La aceptación de Feyerabend parece haber sido pasada por alto por los autores de diversos artículos sobre temas relacionados.]

rado resultados de la inducción, son de hecho *inconsistentes* con sus supuestas premisas inductivas (parciales), como se ha mostrado.

Pero si esto es así, entonces la inducción, en cualquier sentido importante, se derrumba. Y nada más sobre la inducción no probabilística.

En cuanto a un argumento inductivo probabilístico: según el cálculo de probabilidades, si se nos dan varias premisas inductivas consistentes, cualquier conclusión inferida que sea inconsistente con ellas, sólo puede tener, con respecto a las premisas, una probabilidad cero.

3) La teoría de Newton debía mucho, sin duda, a las de Galileo y Kepler; tanto, que el propio Newton consideraba a éstas como premisas inductivas (parciales) de su teoría.

La teoría de Galileo de la caída libre de los cuerpos contenía una constante, *g*, la constante de la aceleración. De la teoría de Newton se sigue que *g* no es una constante sino una variable, dependiente de *a*) la masa del cuerpo que atrae (en el caso de Galileo, la Tierra), y *b*) del cuadrado de la distancia desde el centro de la masa.

Por tanto, la teoría de Galileo es inconsistente con la de Newton.

Naturalmente, en el supuesto de que consideremos sólo los cuerpos en caída libre que estén cerca de la superficie de la Tierra, de forma que todos estén a casi la misma distancia del centro de la Tierra, *podemos explicar por qué* g *parece (equivocadamente) una constante.*

La situación respecto a las leyes de Kepler es muy similar.

Para cada sistema de *dos* cuerpos, uno de los cuales sea muy pesado y el otro tan ligero que su peso sea despreciable, podemos derivar las tres leyes de Kepler a partir de la teoría de Newton y, de ese modo, explicarlas. Pero como Kepler formuló sus leyes para un sistema de muchos cuerpos, que constaba del Sol y varios planetas, esas leyes son, desde el punto de vista de la teoría de Newton, inválidas. Por tanto, no pueden formar ni un sistema completo ni uno parcial de premisas (inductivas o deductivas) para la teoría de Newton.

Esto es todo respecto a la derivación, inductiva o deductiva, de la teoría de Newton a partir de la Kepler o de la de Galileo.

4) Desde luego, constituyó un éxito decisivo para la teoría de Newton el hecho de que pudiera explicar las teorías de Galileo y de Kepler: es decir, el que esas teorías pudieran deducirse de la de Newton bajo ciertas suposiciones simplificadoras (y, estrictamente hablando, falsas).

Pero como las supuestas premisas inductivistas son, estrictamente hablando, inconsistentes con las supuestas conclusiones

inductivistas, es engañoso hablar en este caso de una inferencia inductiva o de una relación probabilística inductiva.

5) Esta clase de situación es típica de la historia de la ciencia. La relación entre la teoría de Newton de la gravitación y la de Einstein es otro ejemplo muy similar y muy importante del mismo caso.

6) Que yo sepa, no se ha dado ninguna respuesta seria a este argumento, hasta ahora, y sobre todo, no la han ofrecido los defensores de ninguna de las teorías probabilísticas de la inducción actualmente en vigor.

16. LAS DIFICULTADES DEL REALISMO METAFISICO. POR UN REALISTA METAFISICO

Sería un error dejar el tema del realismo metafísico sin mencionar, al menos, las dificultades de esta posición. Estas dificultades son graves. A mí me parece plantear problemas insolubles. Y sin embargo, son de un carácter tal que no afectan, en lo más mínimo, a mi fe en el realismo. Están en un plano diferente, por así decirlo, que los problemas y argumentos con los que puedo sustentar mi fe en el realismo. Tal vez sea un plano menos racional, en el que los argumentos se hacen algo vagos y son menos manejables.

Newton se vio impelido a creer, por su teoría de la acción a distancia, que el espacio era el sensorio de Dios. El argumento es, sin duda, algo fantástico; pero hay más en él de lo que parece a primera vista. Porque la dificultad es muy real. Las distancias en el universo son tremendas. La acción a distancia significaría que los efectos gravitatorios estarían, como la Deidad, omnipresentes en todo el universo. Newton, como Einstein, se sentía incapaz de aceptar la acción a distancia como una propiedad de la mecánica de la naturaleza. Sintió su misterio y se lo atribuyó a Dios.

Einstein resolvió este problema, o así lo parece, por medio de su teoría de que las perturbaciones gravitatorias se extienden con la velocidad de la luz. Esta solución es sumamente satisfactoria, sobre todo desde el punto de vista de nuestra discusión en la última sección: indica una posible unificación de las teorías de la luz y de la gravedad y lo hace interpretando la luz y las perturbaciones gravitatorias en términos de propiedades estructurales —propiedades de campos— del universo, de nuestro mundo.

Y sin embargo, seguimos confrontados al problema de Newton. Porque ¿qué pasa con las propiedades estructurales mismas

de nuestro mundo? Son, creemos, las mismas en todo lugar y tiempo. ¿Cómo hemos de entender esto?

Cuando hablamos de las propiedades estructurales de nuestro mundo, hablamos del mundo, metafóricamente al menos, como si fuese una cosa, como un cristal o un globo o quizá una máquina. Pero según la teoría física actual, las propiedades estructurales de un cristal o un globo o una máquina se deben a interacciones entre las partes; y esas interacciones implican velocidades finitas, hasta, e incluida, la velocidad de la luz. Las interacciones mantienen la forma del cristal; determinan la presión del gas en el globo; mantienen montada la máquina.

Pero las propiedades estructurales del mundo que describimos por medio de las leyes de la naturaleza no pueden entenderse así. No pueden, parece, explicarse como debidas a la interacción, puesto que ellas son la base de toda interacción. Son —ésta su característica más profunda— las mismas en todo el mundo, en cualquier lugar y tiempo: son *omnipresentes*. Y así, vuelve a surgir el problema de Newton.

Además, las estructuras de las diversas cosas que hemos mencionado pueden entenderse, o explicarse, en términos de leyes. Pero la estructura del mundo es diferente en el sentido de que es lo que la ley describe en vez de explicar. Esta dificultad, sin embargo, es menos grave. Porque podemos explicar las leyes —y, por tanto, la estructura del mundo— por medio de leyes más profundas: aquí ayuda el haber renunciado a la teoría esencialista de la explicación última. Además éste es precisamente el lugar apropiado para recordarnos a nosotros mismos que no debemos convertirnos en víctimas de nuestra propia metáfora; que «la estructura de nuestro mundo» era sólo una metáfora destinada a ayudarnos a contemplar lo que las leyes describen; que la metáfora tenía que fallar innevitablemente en algún punto; y que es bastante satisfactorio haber encontrado el punto en el que falla.

Todo esto debe tenerse presente. Pero no resuelve la dificultad de Newton. Debemos, creo, aceptar la existencia de leyes de la naturaleza; debemos hacerlo, me temo, como un misterio que se ha hecho aún más impenetrable desde Einstein; porque las propias leyes de la naturaleza que postulan, según Einstein, que ningún efecto puede extenderse con velocidad superior a la de la luz, hacen imposible entender la omnipresente homogeneidad estructural del mundo.

La teoría del universo en expansión puede ayudarnos aquí; pero no si se supone que en un segundo (o en la fracción de un segundo) desde el momento cero, el radio del mundo fue del orden de muchos años-luz. (Parece obvio, al menos a mí me lo

parece, que en ese caso no podría hablarse de «equilibrio térmico».)

Se ha dicho con frecuencia que —por usar las palabras de Wittgenstein— «lo místico no es *cómo* es el mundo, sino *qué* es»[129]. Sin embargo, nuestra discusión muestra que *cómo* es el mundo —que tiene una estructura o que sus regiones enormemente distantes están todas sujetas a las mismas leyes estructurales— parece ser inexplicable en principio y, por tanto, «místico», si queremos usar ese término. Este, en cualquier caso, es el aprieto en el que se encuentra el realista. El idealista puede tener una salida, una explicación a través de la cual puede reducir este misterio al de la existencia misma del mundo. Porque podemos decir con Kant que nuestro intelecto impone sus leyes a la naturaleza; o, en palabras de Wittgenstein, que «sólo son *pensables* las conexiones *legiformes*»[130]. Aunque el realista pueda estar de acuerdo, al menos en parte, con esas concepciones, éstas no le ayudan en lo más mínimo a explicar ni a entender por qué, si tenía que haber un mundo, tenía que ser un mundo pensable, regulado por leyes, un mundo inteligible para algún intelecto; un mundo habitable por vida.

Puede ser útil, quizá, alguna ilustración más sobre la dificultad. Podemos adoptar la terminología de Comte y de Mill, que distinguen dos tipos de leyes o regularidades: regularidades de sucesión y regularidades de coexistencia. Las leyes de sucesión son aquellas leyes «causales» en las que el *tiempo* juega un papel esencial, por ejemplo, las leyes que determinan cambios (tales como las aceleraciones) en el estado de un sistema. Las leyes de coexistencia son, por ejemplo, las leyes que describen las regularidades anatómicas o *estructurales* de un animal, de una molécula o de un átomo.

Ahora bien, las leyes estructurales de coexistencia de los animales, moléculas o incluso átomos, *pueden*, en principio, reducirse a leyes «causales», a aquellas leyes causales según las cuales se producen estas estructuras y se mantienen (relativamente) permanentes durante un tiempo. Parecemos poder entender, al menos en principio, las condiciones de estabilidad de una molécula, por ejemplo, con ayuda de la teoría de la resonancia; esto es, de una interacción causal de las partes constituyentes. Pero no entendemos, hasta el momento, dichas leyes estructurales o leyes de coexistencia, como la constancia absoluta de la carga electrónica o de la masa; o, más en general, la absoluta identidad cualitativa y cuantitativa de las pro-

[129] L. Wittgenstein, *Tractatus Logico-Philosophicus*, 6.44.
[130] *Loc. cit.*, 6.361.

piedades de las partículas elementales. Estas no pueden ser entendidas, al parecer, como debidas a la interacción: según la relatividad general, no puede haber interacción igualadora entre cargas electrónicas simultáneas cuya distancia pueda medirse en años-luz.

Con todo, quizá un día podamos derivar, y explicar, las condiciones de equilibrio de los electrones y otras partículas elementales de una manera que, en apariencia, no sea totalmente diferente de la forma en que explicamos las condiciones de equilibrio de las moléculas, es decir, en términos de «leyes de sucesión» causales, por ejemplo, de ecuaciones de campos que determinen una gama de soluciones discretas. Pero ¿significaría esto que habríamos transcendido el dualismo entre las leyes de sucesión y las leyes de coexistencia reduciendo las últimas a las primeras? De ningún modo; porque, en lugar de preocuparnos por lo inexplicable del hecho —la ley estructural— de que todos los electrones en nuestro vasto mundo parezcan tener, *sin influirse causalmente unos a otros*, absolutamente la misma carga, debemos preocuparnos ahora por el hecho de que todas las partes de nuestro vasto mundo se gobiernen por las mismas leyes que determinan la identidad de las cargas electrónicas. Y que algún día podamos explicar causalmente este hecho estructural —que es claramente el mismo que antes, aunque expresado de un modo diferente— parece obviamente más allá de toda esperanza, puesto que cualquier explicación causal tendría que ser en términos de leyes como aquéllas cuya validez universal querríamos explicar y entender.

Así pues, la aserción de la universalidad o constancia de nuestras leyes causales a través del espacio (y del tiempo) equivale a afirmar una regularidad estructural, una regularidad de coexistencia que parece ser en principio inexplicable causalmente, puesto que no puede explicarse por ninguna ley causal o de sucesión. De hecho, si hablamos de una ley de sucesión, queremos decir que todas las sucesiones, incluidas todas las sucesiones que coexisten, presentan la misma regularidad, la misma estructura. Así, vemos una vez más que la homogeneidad estructural del mundo parece resistirse a una explicación «más profunda»: continúa siendo un misterio.

No creo que pueda resolverse este misterio pensando que el mundo es lo que es por una especie de necesidad lógica. La esperanza de reducir la ciencia natural a lógica me parece al mismo tiempo absurda y repulsiva. Tampoco creo que el idealismo pueda resolver el misterio. El idealismo en sus diversas formas, y sobre todo el convencionalismo y el instrumentalismo, ofrecen soluciones; pero esas soluciones parecen claramente falsas y es preferible enfrentarse a un misterio que tratar

de rehuirlo con soluciones falsas, sobre todo si son fáciles.

Tomemos la ingeniosa solución de Kant: que nuestro intelecto no lee las leyes en el libro abierto de la naturaleza, sino que impone sus propias leyes a la naturaleza. Esto es verdad hasta cierto punto; nuestras teorías son creación nuestra y nunca podemos describir hechos empíricos (o reaccionar ante ellos de otro modo) sin interpretarlos en términos de nuestras teorías (o de nuestras expectativas, tal vez insconscientes). Pero esto no significa, como creía Kant, que las leyes de la naturaleza, tales como la teoría de Newton, sean válidas *a priori* e irrefutables, aunque es cierto que las imponemos sobre los hechos empíricos mismos a los que tendríamos que acudir en busca de refutación. Por el contrario, hemos aprendido de Einstein que nuestro intelecto puede formar, al menos tentativamente, teorías alternativas; que pueden reinterpretar los hechos alternativamente en términos de cada una de esas nuevas teorías; que, en la competición entre esas teorías, podemos decir libremente, sondeando su profundidad y pensando el resultado de nuestras críticas, incluyendo nuestras contrastaciones; y que sólo de este modo podemos tener la esperanza de acercarnos a la verdad.

La epistemología de Kant ha sido refutada también, creo, por el hecho mismo de que tiene sumo éxito, pero *sólo dentro de límites muy extrechos.* Kant creía que estábamos dotados de un aparato mental, con un sistema digestivo psicológico y quizá fisiológico, que nos permitía digerir los estímulos que llegaban a nuestros sentidos desde el mundo exterior y que, al digerirlos, asimilarlos y absorberlos, imponíamos nuestras características estructurales sobre ellos; es decir, es la marca de nuestra mente sobre ellos la que da lugar a verdades válidas *a priori* sobre el mundo. Ahora es innegable que existen tales verdades genéticamente *a priori.* Pero son curiosamente poco importantes. Porque *no* estamos obligados a interpretar el mundo de las cosas según ellas; por el contrario, nos percatamos fácilmente de su carácter subjetivo y las tratamos como se merecen. Ejemplos excelentes de este mecanismo kantiano nos los proporcionan ciertas ilusiones ópticas regulares e inescapables. Otros nos lo ofrece el orden de los colores; es decir, el hecho que experimentemos que el rojo es más similar al anaranjado, amarillo, morado y azul que al verde; el amarillo más similar al verde y al rojo que al azul, etc. Estas son verdades *a priori*: puede ser una verdad *a priori* que percibimos los colores, que no somos daltónicos; puede ser también un hecho empírico que nuestro mecanismo para la percepción del color esté basado en un componente rojo-verde y amarillo-azul en vez de estarlo en uno rojo-azul y amarillo-verde (estoy usando aquí los nombres de

los colores para las longitudes de onda y no para las percepciones). Pero una vez que vemos los colores con ayuda de este mecanismo, los ordenamos necesaria e intrísecamente según esas relaciones de similitud y disimilitud.

La explicación de todo esto es, sin duda, el mecanismo digestivo psico-fisiológico kantiano del que estamos dotados. Nuestra fisiología explica esas similaridades. Vemos con ayuda de un componente rojo-verde o amarillo-azul; y si falta uno de los dos, somos daltónicos, o daltónicos respecto al rojo y el verde, o daltónicos respecto al amarillo y el azul. El rojo no puede, fisiológicamente, fundirse con el verde, ni el azul con el amarillo. Así pues, se convierten en opuestos. Los otros colores pueden fundirse, y uno puede, matizándose, convertirse en otro, porque los dos componentes pueden ser estimulados independientemente al mismo tiempo.

Kant tenía razón en esto. Pero la cuestión resulta trivial por comparación. Estas relaciones válidas *a priori* entre los colores no llevan a nadie a imponer equivocadamente las leyes correspondientes a la naturaleza, al mundo de las cosas con color. Ni creemos que las *cosas* rojas tengan, desde un punto de vista físico, mayor afinidad con las amarillas o las azules que con las verdes; ni pensamos que haya leyes de la naturaleza por descubir en esto, excepto, naturalmente, leyes de nuestra propia psicología y fisiología. Pero según Kant, nosotros imponemos estas leyes de nuestro aparato digestivo mental sobre el mundo, en el sentido de que están abocadas a convertirse en leyes «objetivas» de las cosas que percibimos.

Las ilusiones ópticas pertenecen a la misma categoría. Y podría conjeturarse que nuestra creencia en las *cosas* reales está igualmente fundada fisiológicamente. Pero en este caso, los mecanismos fisiológicos y las creencias que emanan de ellos (ambos son, podemos conjeturar, los resultados de una larga evolución y adaptación) parecen resistir a las críticas y vencer en la competición con teorías alternativas. Y cuando nos engañan efectivamente, como en el cine —sobre todo en los dibujos animados— no llevan a los adultos a afirmar en serio que estamos ante un mundo de cosas. Así pues, no somos (como pensó Kant y también Hume) las víctimas de nuestra «naturaleza humana» o de nuestro aparato digestivo mental, de nuestra psicología o fisiología. No somos prisioneros de nuestras mentes para siempre. Podemos aprender a criticarnos a nosotros mismos y, de ese modo, a transcendernos a nosotros mismos. Tenemos limitaciones; pero somos más libres de lo que Kant pensaba.

Consideraciones similares pueden aplicarse a otras formas de esclavitud humana, tales como la tendencia a aceptar las

valoraciones, creencias y dogmas de nuestro grupo social. Esta tendencia es fuerte y puede tener también una base fisiológica. Pero podemos escapar de ella. Para ello podemos necesitar al principio, quizá, el estímulo de un choque cultural o de una conversión. Pero más adelante podemos desarrollar el hábito de inquirir, de la discusión crítica, racional. La discusión racional y el pensamiento crítico no son como los sistemas primitivos de interpretar el mundo; no son un marco al que estamos atados y amarrados. Por el contrario, son los medios para escapar de la prisión, de liberarnos.

Hume enseñó que las creencias humanas fundamentales no eran más que hábitos adquiridos de forma irracional que los hombres no pueden trascender, sino que están obligados a obedecerlos. Kant, en un sentido, aceptó esta idea pesimista aunque su «revolución copernicana» le dio un cariz optimista. El enseñó que, puesto que el mundo objetivo de la experiencia era un mundo formado por nuestro intelecto experimentante (que representaba este papel incluso en el plano de la percepción), nuestras creencias sobre este mundo podían ser objetivamente verdaderas y racionales: aunque pudieran describirse como hábitos, no eran sólo hábitos, puesto que nuestra razón tomaba parte en su formación en la misma medida que nuestros sentidos y nuestras asociaciones. Todo desarrollo ulterior de esta línea de pensamiento debe tomar nota del hecho de que el *aumento del conocimiento* consiste fundamentalmente en la revisión crítica de nuestras creencias, un hecho que establece que no estamos ligados a nuestras creencias fundamentales en el grado en que pensaban Hume y Kant. También establece que tanto Hume como Kant tenían razón en parte. Hume tenía razón en ser escéptico en cuanto a la validez de nuestras creencias: esto lo muestra el hecho de que hayamos transcendido algunas de ellas (lo que sugiere la posibilidad de que transcendamos otras). Pero, a pesar del gran crítico que era, se equivocaba al pasar por alto, precisamente, la importancia de nuestra capacidad de transcender nuestras creencias criticándolas. Y Kant tenía razón al señalar, frente a Hume, que el razonamiento estaba implicado en la formación de nuestras creencias, incluso de nuestras creencias habituales. Tenía razón, además, al señalar la importancia del aumento de nuestro conocimiento y al enseñar que todo aumento del conocimiento necesita un marco teórico que debe precederlo. Pero se equivocaba al pensar que ese marco, a su vez, no podía transcenderse de ningún modo, y que era, por tanto, *válido a priori*. Continuando nuestra historia, podemos decir quizá que Hegel tenía razón al señalar (si podemos interpretar así sus oscuras enseñanzas) que el marco también estaba sujeto a aumento y que podía ser trans-

cendido. Pero se equivocaba al sugerir (de nuevo, si podemos interpretar así sus enseñanzas) que la verdad es esencialmente relativa respecto a algún marco y que no es *nuestra* crítica activa, *nuestro* descubrimiento de una contradicción, de una refutación, lo que nos fuerza a cambiar nuestras ideas o creencias, sino que esas ideas se transcienden a sí mismas, de modo que dependemos de las ideas en evolución, en vez de ellas de nosotros, de nuestra crítica racional. Esto hace depender a nuestra crítica del marco históricamente heredado y nos lleva, por tanto, al relativismo, al relativismo histórico.

Esta filosofía de la esclavitud humana ha ejercido una extraña fascinación sobre la teoría post-kantiana del conocimiento. (También ha jugado una parte importante en la decadencia del racionalismo y el liberalismo)[131]. Todo lo que hizo Kant por corregir a Hume, lo deshizo la doctrina de que nuestros hábitos de creencias —nuestras valoraciones, actitudes, dogmas y, por tanto, nuestro mundo de experiencia— dependen de nuestra época histórica o de nuestro grupo social (Hegel, Marx). Esta doctrina, que se ha puesto tan de moda en las manos de la moderna ciencia social es cierta, naturalmente, mientras no le atribuyamos ninguna significación epistemológica particularmente profunda. Es cierto que dependemos de nuestra educación, nuestras creencias, nuestro conocimiento, nuestras expectativas. Pero también es cierto que no dependemos totalmente de ellos. Es indudable que sólo podemos liberarnos lenta y parcialmente de estas cadenas. Pero no hay límite natural para este proceso de liberación para el aumento del conocimiento. Es posible, desde luego, negar que vayamos a poder romper algún día nuestros grilletes intelectuales: es posible afirmar que nos engañamos a nosotros mismos si creemos que estamos menos atados por el marco de nuestros prejuicios que las gentes de hace dos mil años, afirmar que el marco, la moda, ha cambiado, pero no su poder sobre nosotros. Aunque esta concepción, tan en boga, de la cuestión puede ser irrefutable, es simplemente falsa. Desde el Renacimiento ha habido un sorprendente aumento de la actitud crítica.

Una forma especial de esta filosofía de la esclavitud humana es el relativismo lingüístico, una concepción que ha sido vigorosamente expuesta por Benjamín Lee Whorf[132]. En nuestro contexto actual puede formularse como la concepción de que

[131] El tema se desarrolla, aunque principalmente respecto a la filosofía pre-kantiana, en mi artículo «Sobre las fuentes del conocimiento y de la ignorancia», *Proceedings of the British Academy* 46, 1960, pp. 39-71; reimpreso en *El desarrollo del conocimiento científico (C. & R.)*.

[132] B. L. Whorf, *Language, Thought and Reality*, 1956.

nuestros lenguajes humanos pueden incorporar (o no incorporar) en sus estructuras creencias, teorías y expectativas en tal grado que no podemos romper esos grilletes ideológicos mediante el recurso a la crítica, puesto que la crítica tiene que usar siempre el lenguaje. Según esta formulación, el uso del plural «lenguajes» puede indicar cómo librarse de esas ataduras: no son tan fuertes como podría creerse, porque es posible que los hombres se liberen unos a otros criticando un lenguaje o un sistema de creencias por medio de otro (choque cultural). No hay razón alguna para pensar, como lo hacen algunos, que Whorf ni ningún otro, haya mostrado lo inconmensurable de los conjuntos de creencias (o que todas las aserciones sean relativas a diferentes conjuntos irreductibles de creencias fundamentales). Otros se han visto llevados a pensar, por los fascinantes análisis de Whorf, que, como todos los lenguajes tienen algo en común, tendrán un conjunto común de creencias que debe ser imposible detectar por ese método de crítica mutua que depende esencialmente de la divergencia lingüística. Hay que admitir que esto es verdad hasta cierto punto; y puede expresarse de modo más sencillo diciendo que siempre albergaremos prejuicios de los que no seamos conscientes. Pero esto no significa que no podamos detectar, a veces, por uno u otro método (o sin método alguno), algunos de nuestros prejuicios y deshacernos de ellos por medio de la crítica. Tampoco significa que este proceso, quizá lento pero continuo, de liberación intelectual no pueda apresurarse con la práctica del pensamiento crítico y de la discusión racional.

La discusión racional no puede practicarse, sin embargo, como un mero juego para matar el tiempo. No puede existir sin problemas reales, sin la búsqueda de la verdad objetiva, sin una tarea de descubrimiento que nos propongamos nosotros mismos: sin una realidad que descubrir, una realidad que explicar por medio de *leyes estructurales universales.*

Así, volvemos a nuestro problema, y al de Newton. El idealismo ofrece una salida fácil, pero ni siquiera en la forma que le dio Kant ofrece una solución convincente. En cualquier caso, nosotros los realistas tenemos que vivir con la dificultad. Pero debemos enfrentarnos a ella[133].

[133] Cuando escribí esta frase estaba convencido de que «el problema de Newton», como lo llamé, era insoluble, o que, en el mejor de los casos, tenía una solución religiosa, insatisfactoria, algo en la línea del propio Newton. Yo personalmente no tenía la esperanza de resolverlo ni intención de abordarlo siquiera. Por tanto, fue una completa sorpresa para mí cuando más tarde, mientras trataba de reinterpretar la relatividad general en el sentido del indeterminismo, tropecé con lo que parecía la solución de una parte —una parte mínima— del problema de Newton. Como consecuencia, ya no estoy conven-

Con esto termino la discusión metafísica que comencé en la sección 7, animado por un pasaje de Einstein que escogí como lema; y con ella, concluyo también la «cuarta fase» —la fase metafísica— del problema de la inducción.

cido de que sea insoluble. (Véase también *El universo abierto* [*O. U.*], tomo II de este *Post Scriptum*, sección 19; y *Teoría cuántica y el cisma en física*, tomo III del *Post Scriptum*, sección 27.)

CAPITULO II

LA DEMARCACION

17. LA SIGNIFICACION DEL PROBLEMA DE DEMARCACION

Después de presentar de nuevo mi solución al problema de la inducción, he tratado de seguir sus ramificaciones hasta su fase metafísica, como yo la llamo, mucho más allá de su esfera original. Con todo, mi exploración de las ramificaciones del problema de la inducción quedaría incompleta si descuidase *el problema de la demarcación entre la ciencia y la metafísica*. En realidad, hay una pregunta que la gente me hace en cuanto se da cuenta de que realmente no creo en la inducción y de que ni siquiera creo que la inducción tenga un papel significativo en las ciencias. Es ésta: si abandona usted la inducción, *¿cómo puede distinguir las teorías de las ciencias empíricas de las especulaciones pseudo-científicas, no científicas o metafísicas?*

Este es el problema de la demarcación. Ha de resolverse, sugiero, aceptando la contrastabilidad o refutabilidad o falsabilidad como la característica distintiva de las teorías científicas. Con la formulación dada, apenas es posible calibrar su significación. A primera vista puede incluso parecer la pregunta de un pedante más que un problema de interés real. Porque ¿qué hay en un nombre o en una distinción o en una clasificación o en una demarcación? Si estamos ansiosos por saber, si nuestro objetivo es el de aprender sobre el mundo, no nos preocupamos mucho por los compartimentos o departamentos a los que pueda ser destinado nuestro conocimiento en perspectiva. Como dije en la Introducción, las disciplinas y otras divisiones del saber son ficticias y muy engañosas, por muy convenientes que sean como unidades administrativas. En lo que respecta a la ciencia y a la metafísica, yo no creo, ciertamente, en una demarcación precisa. La ciencia, en todas las épocas, ha sido profundamente influida por las ideas metafísicas; ciertas ideas y problemas metafísicos (tales como el problema del cambio, o el programa cartesiano de explicar todo cambio por una acción a distancias decrecientes) han dominado el desarrollo de la ciencia durante siglos, como ideas reguladoras; mientras que otras (tales como el atomismo, otro intento de resolver el problema

del cambio)[1] se han transformado gradualmente en teorías científicas. Naturalmente, ha habido evoluciones en el sentido opuesto, también: como les gusta decir a algunos positivistas, puede mostrarse que un número considerable de doctrinas metafísicas son vestigios de doctrinas obsoletas de la ciencia.

Podemos ilustrar esto con ayuda de la historia del propio positivismo. Puede decirse que el positivismo y el fenomenalismo de Mach fueron, en su origen, una teoría científica respetable, concebida para explicar la falta de éxito del atomismo y de otras teorías sobre la estructura de la materia por medio de la hipótesis de que simplemente no existía el ente físico llamado materia o «sustancia». Mach podía señalar al éxito de la física fenomenalista —sobre todo de la termodinámica fenomenalista— y a las dificultades lógicas fundamentales con que se encontraba Boltzmann en sus intentos de explicar la segunda ley en términos de una estructura atómica o molecular. La solución propuesta por Mach suponía que esos problemas y otros relativos a la «sustancia» o «materia» eran pseudoproblemas; incluidos, naturalmente, todos los problemas relativos a «la estructura de la materia». Pero debido a la obra de Einstein de 1905, sobre el movimiento browniano, quedó establecida toda la significación física de las teorías de Maxwell y Boltzmann. El movimiento browniano logró, por medio de la interpretación que le dio Einstein, la posición de un experimento crucial. Y, como señaló el propio Einstein, la existencia del movimiento browniano refutaba la versión fenomenalista de la segunda ley de la termodinámica[2]. Con esto se mostró que el problema de la estructura atómica de la materia era un auténtico problema físico. Así, después de 1905, el positivismo y el fenomenalismo de Mach se hicieron crecientemente metafísicos en uno de los sentidos favoritos de los positivistas: se convirtió en un trozo de la física obsoleta que los científicos *qua* científicos habían abandonado, pero que continuó subsistiendo entre los filósofos y entre los científicos, cuando se volvían filósofos

[1] Véase la sección 23 *infra* (texto anterior a la nota 37).

[2] Este es, pues, otro caso de una teoría que corrige su propia base observacional; véase la sección 15, *supra* sobre todo el texto después de la nota 125, y también mi artículo «Irreversibility, or Entropy since 1905», *Brit. Journ. Philos. Science* 8, 1957. [Véase también Popper, «The Arrow of Time», *Nature*, 17 de marzo, 1956, p. 538; «Irreversibility and Mechanics», *Nature*, 18 de agosto, 1956, pp. 381-2; «Irreversible Processes in Phisical Theory», *Nature*, 22 de junio, 1957, pp. 1296-7; «Irreversible Processes in Physical Theory», *Nature*, 8 de febrero, 1958, pp. 402-3; «Time's Arrow and Entropy», *Nature*, 17 de julio, 1965, pp. 233-4; «Time's Arrow and Feeding on Negentropy», *Nature*, 21 de enero, 1967, p. 320; «Structural Information and the Arrow of Time», *Nature*, 15 de abril, 1967, p. 322; y *Búsqueda sin término. (u. Q.), op. cit.*, sección 35, *Ed.*]

o apologistas, como hacen a veces cuando sus teorías se ven en apuros. (Véase también la sección *113, *infra*, es decir, la sección 21 de *Teoría Cuántica y el cisma en física*, tomo III del *Post Scriptum*.)

Como lo muestran esos ejemplos, no puede haber una demarcación exacta entre la ciencia y la metafísica; y la importancia de la demarcación, si la tiene, no debe sobreestimarse. A pesar de esto, yo mantengo que el problema de la demarcación es sumamente significativo. Lo es, no porque tenga un mérito intrínseco el clasificar teorías, sino porque una serie de problemas auténticos e importantes están estrechamente vinculados con él, en realidad, todos los problemas principales de la lógica de la ciencia.

Al principio de esta sección hice alusión a uno de esos vínculos: a la concepción de que el método inductivo nos proporciona un criterio de demarcación. Otro mencionado anteriormente —el problema de la discutibilidad de las hipótesis científicas y, por tanto, de su racionalidad— está vinculado, desde luego, con el problema de su contrastabilidad. Podemos considerar la contrastabilidad como un cierto tipo de discutibilidad: discutibilidad por medio de argumentos *empíricos*, argumentos que recurren a la observación y al experimento. Un tercer vínculo con el problema de la inducción se pone de manifiesto en el modo en que he diferenciado la cuarta fase o fase metafísica del problema de la inducción de sus tres fases lógicas o metodológicas. En su cuarta fase —es decir, como el problema de si existen o no leyes naturales verdaderas— el problema adoptó un carácter sorprendentemente diferente del de las fases anteriores y esta diferencia requería una dilucidación urgente. Para esta dilucidación, el carácter existencial del problema nos proporcionaba la clave: los enunciados puramente existenciales son irrefutables empíricamente. Si hay que discutirlos en absoluto, debemos tener siempre presente su irrefutabilidad empírica. He tratado de establecer el hecho de que los enunciados y problemas metafísicos pueden ser de todos modos discutibles (aunque no concluyentemente), por el simple mecanismo de discutir sobre ellos.

Naturalmente, el problema de demarcación está también estrechamente relacionado, histórica y lógicamente, a lo que llamé al principio de la sección 2 el problema central de la filosofía del conocimiento. Pues el problema de cómo juzgar o decidir entre teorías o creencias rivales lleva, como dije allí, al problema de decidir si es posible o imposible justificar racionalmente una teoría; y esto, a su vez, lleva al problema de diferenciar o deslindar las teorías racionales de las creencias irracionales; un problema que a menudo se identifica (quizá un

poco precipitadamente) con el problema de diferenciar o deslindar las teorías empíricas o «científicas» de las «metafísicas».

Así el problema de demarcación es más que una cuestión de clasificar las teorías para poder llamarlas «científicas» o «metafísicas». En realidad da acceso a algunos de los problemas más fundamentales de la teoría del conocimiento y, por tanto, de la filosofía.

Pero el problema de demarcación tiene también una importancia práctica considerable. Tropecé con este problema y con su solución varios años antes de interesarme por el problema de la inducción y antes de percibir estos vínculos entre el problema de la inducción y el de la demarcación a los que acabo de referirme. Esto ocurrió en 1919, cuando empecé a sospechar de las diversas teorías psicológicas y políticas que reclamaban la categoría de ciencias empíricas, sobre todo el «psicoanálisis» de Freud, la «psicología individual» de Adler y la «interpretación materialista de la historia» de Marx[3]. Todas esas teorías se sostenían de un modo *acrítico*, me parecía. Se introdujo gran número de argumentos en su apoyo. Pero la crítica y los contra-argumentos se consideraron como hostiles, como síntomas de un rechazo premeditado a admitir la verdad manifiesta; y, por tanto, se les opuso hostilidad en lugar de argumentos.

Lo que yo encontraba muy llamativo, y muy peligroso, en esas teorías, era la afirmación de que estaban «verificadas» o «confirmadas» por un flujo incesante de evidencia en forma de observaciones. Y, realmente, una vez que se le abrían a uno los ojos, podía ver en todas partes ejemplos que las verificaban. Un marxista no podía mirar un periódico sin encontrar evidencia verificadora de la lucha de clases en cada página, desde los artículos de fondo hasta los anuncios; y también la encontraba, y muy especialmente en lo que el periódico no decía. Y un psicoanalista, sea freudiano o adleriano, dirá, con toda seguridad, que encuentra verificadas sus teorías cada día, incluso cada hora, por sus observaciones clínicas.

Pero ¿eran contrastables esas teorías? ¿Estaban esos análisis realmente mejor contrastados que, por ejemplo, los frecuentemente «verificados» horóscopos de los astrólogos? ¿Qué suceso concebible podía falsarlos a los ojos de sus partidarios? ¿No era todo suceso concebible una «verificación»? Era precisamente ese hecho —el hecho de que siempre encajasen, de que siempre resultasen verificados— lo que impresionaba a sus partidarios. Empecé a darme cuenta de que esta fuerza aparente era en

[3] He contado esta historia desde un ángulo algo diferente en mi artículo «Philosophy of Science: A Personal Report». Véase también *Búsqueda sin término (U. Q.)*, sección 8.

realidad una debilidad y que todas esas «verificaciones» eran demasiado fáciles para contar como argumentos.

El método de buscar verificaciones me pareció incorrecto; en realidad me pareció el método típico de una pseudo-ciencia. Advertí la necesidad de diferenciar este método lo más posible del otro —el método de contrastar una teoría tan rigurosamente como es posible—, es decir, *del método de la crítica, del método de buscar ejemplos falsadores.*

El método de buscar verificaciones no sólo era acrítico: fomentaba además una actitud acrítica tanto en el expositor como en el lector. Amenazaba, pues, con destruir la actitud racional, el argumento crítico.

Freud fue, con mucho, el más lúcido y persuasivo de los expositores de las teorías a las que estoy refiriéndome. Pero ¿cuál era su método de argumentar? Daba ejemplos; los analizaba; y mostraba que encajaban en su teoría, o que su teoría podía describirse como una generalización de los casos analizados. Algunas veces pedía a sus lectores que pospusiesen sus críticas, e indicaba que respondería a todas las críticas en una ocasión posterior. Pero cuando me fijé más atentamente en una serie de casos importantes, encontré que las respuestas no llegaban. Sin embargo, por extraño que parezca, muchos lectores se consideraban satisfechos.

Para mostrar que éstas no son meras aserciones ni acusaciones vacuas, las probaré detalladamente por medio de un análisis de la discusión de Freud de la tesis fundamental de su gran libro, *La interpretación de los sueños,* considerado correctamente por él y por otros como su obra más importante. *¿Fue crítico su enfoque?*

18. UN CASO DE VERIFICACIONISMO

> *Si un paciente que, de otro modo, es muy inteligente, rechaza una sugerencia con argumentos no demasiado inteligentes, entonces la imperfección de su lógica es evidencia de la existencia de un... fuerte motivo para su rechazo.*
>
> SIGMUND FREUD

El propósito de esta sección es mostrar, analizando un caso famoso, que el problema de demarción no es simplemente un problema de clasificar las teorías en científicas y no científicas, sino que su solución se necesita urgentemente para poder realizar una evaluación crítica de las teorías científicas o de las teo-

rías supuestamente científicas. Para este propósito, he seleccionado la gran obra de Freud, *La interpretación de los sueños*, por dos razones. La primera es que mis intentos de analizar sus argumentos tuvieron un papel importante en el desarrollo de mis concepciones sobre la demarcación[4]. La segunda es que, a pesar de defectos graves, algunos de los cuales trataré de exponer aquí, contiene, fuera de toda duda razonable, un gran descubrimiento. Yo, al menos, estoy convencido de que existe un mundo del inconsciente y de que los análisis de los sueños de Freud en su libro son fundamentalmente correctos, aunque sin duda incompletos (como el mismo deja claro) y, necesariamente, algo sesgados. Digo «necesariamente» porque incluso la observación «pura» no es nunca neutral, es necesariamente el resultado de una interpretación. (Las observaciones siempre se recogen, ordenan, descifran y pesan a la luz de nuestras teorías. En parte por esa razón, nuestras observaciones tienden a apoyar nuestras teorías. Este apoyo de poco o ningún valor a menos que adoptemos conscientemente una actitud crítica y busquemos refutaciones a nuestras teorías, en vez de «verificaciones».) Lo que sea válido incluso para las observaciones más imparciales, lo será también para la interpretación de los sueños.

Lo que me propongo hacer en esta sección es analizar el modo de argüir de Freud en apoyo de su tesis central en *La interpretación de los sueños*.

El principal objetivo de Freud en este libro es del de «demostrar que, en su naturaleza esencial, los sueños representan satisfacciones de deseos»[5]. Freud es consciente, natural-

[4] Otra teoría que tuvo un papel similar (véase «Philosophy of Science: A Personal Report») fue el marxismo. (Véase *Búsqueda sin término* [*U. Q.*], sección 8; pero, mientras que he tratado sobre marxismo con gran detalle en *La sociedad abierta y sus enemigos* [*O. S.*] y sobre el historicismo en general en *La miseria del historicismo* [*Poverty*], no había publicado nunca un análisis detallado del método que tiene Freud de abordar los ejemplos falsadores y las sugerencias críticas.)

[5] *Cf.* Sigmund Freud, *The Interpretation of Dreams* (publicada por primera vez en 1899), traducida [al inglés] y editada por James Strachey, 1954, p. 127. Véanse también pp. 119, 121. En lo que sigue, traduzco directamente, en uno o dos lugares, de *Gesammelte Schriften*, II y III, 1925. Puedo añadir aquí que un análisis de *Introductory Lectures* (1916-1917) hubiera llevado a los mismos resultados. (*Cf.* sobre todo la décimo-cuarta conferencia.) Para una crítica psicoanalítica y el reestablecimiento de la tesis principal de Freud, véase J. O. Wisdom, «A Hypothesis to Explain Trauma Re-enactment Dreams», *Intern J. of Psycho-Anal*, 30, 1949, pp. 13 ss. Los pasajes de Freud que se citan ahí completan los que yo he citado aquí. Compárese sobre todo la referencia de Wisdom en pp. 13 y 15 (notas 2 y 8) a Freud, *New Introductory Lectures*, 1937, pp. 43-44, donde Freud introdujo, para explicar los sueños de representación de traumas, el concepto de «*intento*» de satisfacción de deseos; «pero no consideró que con esto», escribe Wisdom en p. 15, «decía nada esencialmente

mente, de que hay una objeción obvia a esta teoría: la existencia de pesadillas y de *sueños de angustia*; sin embargo, rechaza esta objeción. «En efecto, parece» —escribe Freud en un pasaje en el que formula lo que va a ser nuestro problema principal aquí— «que los sueños de angustia hagan imposible aseverar al proposición general (basada en los ejemplos citados en el capítulo precedente) según la cual los sueños son satisfacciones de deseos; realmente los sueños de angustia parecen calificar a esa proposición de absurda. No obstante, *no hay gran dificultad* en responder a esa objeción»[6]. El método de responderla, explica[7], consiste en mostrar que lo que en *apariencia* (en su «contenido *manifiesto*») parece ser un sueño de angustia, en *realidad* (en su «contenido *latente*») es la satisfacción de un deseo. Esto le lleva a Freud a una ligera «modificación» de su tesis principal en relación con «la naturaleza esencial de los sueños», modificación que formula como sigue: «*un sueño es la satisfacción (enmascarada) de un deseo (suprimido o reprimido)*»[8].

Freud reafirma repetidamente su programa de revelar el contenido latente de todo sueño de angustia como satisfacción de un deseo. Así, el programa se reafirma, por ejemplo, en la p. 550 y más plenamente en la 557, donde leemos: «Así, *no hay dificultad* para ver que los sueños desagradables y los sueños de angustia son una satisfacción de deseos en el sentido de nuestra teoría, en la misma medida en que los son los sueños de franca satisfacción»[9]. *Sin embargo, Freud nunca lleva a cabo su programa; y al final, renuncia por completo a él,* aunque sin decirlo explícitamente. La evidencia en que se basa esta aserción es la siguiente:

Freud, bastante al principio de su libro (p. 157), empieza a discutir «los muy frecuentes sueños que parece estar en contradicción con mi teoría»[10]; y muy pronto tenemos indicios de que

---

nuevo». Compárese también la referencia de Wisdom en p. 14 (notas 4 y 5) a la explicación de Freud de los «sueños dolorosos» por la satisfacción de «*deseos de castigo*», como se indica en *Introductory Lectures*, 1943, p. 185 y *Beyod the Pleasure Principale*, 1922, p. 38. Véase también la nota 24 *infra*.

[6] *Op. cit.*, p. 135. (Los subrayados son míos.)

[7] *Loc. cit.*

[8] *Op. cit.*, p. 160. La tesis principal de Freud está estrechamente relacionada con otra fundamental (*cf.* pp. 123 y 133 ss.): que es «la función» de un sueño o, en cualquier caso, su función «normal», ser *guardián* de la dormición contra las perturbaciones; a pesar de que a veces el sueño también puede tener que «aparecer en el papel de *perturbador* de la dormición» (p. 580).

[9] *Op. cit.* p. 537. El subrayado es mío. Compárese esta cita con el texto de mi nota 6 *supra*. Sin embargo, Freud ha hablado de los sueños de angustia en términos muy diferentes y menos confiados antes de este pasaje: por ejemplo, en pp. 161 y 236 (véase *infra*).

[10] *Op. cit.* p. 157. La discusión de los «sueños desagradables» se continúa en pp. 556 ss.; véase la cita de mi nota precedente. Por cierto, estoy dispuesto

el programa de reducir los sueños de angustia a sueños de satisfacción de deseos puede quedar como un sueño de deseo insatisfecho; porque en p. 161 nos enteramos de que *en los sueños de angustia, hay que separar la angustia del sueño* a cuyo contenido sólo «está superficialmente ligada» (véase también p. 162 y la nota del editor en ella). En p. 236, nos enteramos de que la angustia «puede ser psico-neurótica... *Donde esto ocurre... nos acercamos al límite en el que falla el propósito de los sueños de satisfacer deseos*». (La cursiva es mía; véase también el final de p. 487.) De modo que hay un límite, después de todo. En p. 580, el propio Freud se hace consciente de que hasta el momento no ha hecho más que evadir la cuestión de reducir los sueños de angustia a sueños de satisfacción de deseos: «Me refiero, naturalmente», escribe, «a la cuestión del sueño de angustia; y para no confirmar la impresión de que estoy tratando de evadir la evidencia de este testigo principal contra la teoría de la satisfacción de deseos siempre que me encuentro frente a ella, daré ahora, al menos, ciertos indicios de una explicación del sueño de angustia»[11]. Pero los indicios son insatisfactorios; al menos, no satisfacen a Freud. Porque, tras dos páginas en las que no aparece nada nuevo que pueda iluminarnos respecto a nuestro problema, sino solamente una repetición de la vieja aserción de que «no hay ya nada contradictorio en la noción de que un proceso físico que produce ansiedad pueda ser, no obstante, la satisfacción de un deseo», Freud renuncia por completo al intento. Finalmente nos dice, en p. 582, que todo el asunto de los sueños de angustia queda definitivamente «fuera del marco psicológico de la formación de los sueños. Si no fuera por el hecho de que nuestro tema [la teoría de los sueños] está relacionado con la angustia por el solo factor de la liberación del inconsciente durante el sueño, *me sería posible omitir toda discusión de los sueños de angustia y evitar la necesidad de penetrar, en estas páginas, en todas las oscuridades que los rodean*»[12]. En 1911, pero no en las ediciones posteriores, Freud resumió su elaborado, aunque sólo implícito y aparentemente incosciente, repudio de su programa en una sola oración: «La angustia en los sueños, insistió, es un problema de angustia y no un problema de sueños»[13].

---

a creer la hipótesis de Freud (p. 157) de que algunos de sus pacientes satisfacían, en sus sueños, el *deseo* de refutar la teoría de Freud. Y sin embargo, con esta hipótesis nos acercamos peligrosamente a una estratagema convencionalista (*cf. L. I. C.* sección 20) como trataré de argüir.

[11] *Ges. Schriften*, II, 1925, p. 497; correspondiente a p. 580 (las cinco últimas líneas antes del párrrafo nuevo) de la edición de Strachey.

[12] Edición de Strachey, p. 582; los subrayados son míos.

[13] *Loc. cit.*, nota 2; los subrayados son míos. Un enunciado bastante ine-

En las cuatro páginas siguientes, Freud discute y analiza en parte tres sueños de angustia. Su propósito no es ya demostrar que son cumplimiento de deseos, sino simplemente apoyar su aserción de que «la angustia neurótica tiene un origen sexual» (p. 582). Esto, claro está, implica la concepción de que la angustia está vinculada a ciertos *deseos*. Pero no justifica, en absoluto, la inferencia de que todo sueño de angustia tenga que tener el carácter de una *satisfacción* de deseos. (Esta inferencia equivocada parece haber sido hecha por algunos de los lectores de Freud; pero hay que advertir que el propio Freud sugiere solamente que el primero de los tres sueños puede haber sido, *en parte*, la *satisfacción* de un deseo y no sugiere nada por el estilo sobre el segundo y el tercero de los sueños.)

La razón por la que Freud no realiza su programa original de mostrar (por medio de análisis detallados como los que suele dar) que todos los sueños de angustia son satisfacciones de deseos es, claramente, que al final él mismo no lo cree así. De este modo, el sueño de angustia se convierte en un problema de angustia: ahora «forma parte de la psicología de las neurosis» (p. 582) más que de la teoría de los sueños, es decir, de la teoría de la satisfacción de los deseos. Yo seré el último en criticar este cambio de opinión. Pero el cambio no supone una corrección consciente ni la admisión de un error. Por el contrario, nueve años después de que se escribieran esos pasajes, Freud añadió a la página (135) en la que había introducido por primera vez su programa de reducir los sueños de ansiedad a sueños de satisfacción de deseos, un áspero reproche a «los lectores y críticos de este libro». Les acusa de no estar de acuerdo con su tesis de que todos los sueños, incluso los de ansiedad, son satisfacciones de deseos y de no entender su programa (abandonado años antes, aunque sólo al final del libro) según el cual «los sueños de angustia, una vez interpretados, pueden resultar ser satisfacciones de deseos» (p. 135). «Es casi imposible dar crédito a la obstinación», escribe Freud, «con la que los lectores y críticos de este libro cierran sus ojos a esta consideración y a la fundamental distinción entre el contenido manifiesto y el contenido latente de los sueños»[14].

---

quívoco (que, sin embargo, no contiene el término «angustia», sino el de «neurosis traumática» en su lugar), respecto al hecho de que algunos sueños de angustia no son satisfacciones de deseos, pero que «son las únicas excepciones *auténticas*», se encuentra en la primera oración de la sección IX del artículo de Freud de 1923, al que se refiere la nota 16 *infra; cf.* también las cuatro últimas páginas del primer capítulo de *New Introductory Lectures*, 1933, especialmente la observación «no tendré el recurso de decir que la excepción confirma la regla...».

[14] *Ges. Schriften*, III, p. 25; correspondiente a la nota 2, p. 135 de la tra-

Ahora bien, lo que quiero tratar no es tanto que, de hecho, no fueron los lectores y los críticos los obstinados; que los lectores y los críticos no podían dejar de ver el problema de los sueños de angustia; y que tenían toda la razón si quedaban insatisfechos al oírse decir primero que la reducción de los sueños de angustia a sueños de satisfacción de deseos no presentaba «gran dificultad» (pp. 135 y 557) y al encontrarse al final (p. 582) con que esta reducción ni se intentaba siquiera, sino que se rechazaba por «no ser un problema de sueños». Lo que quiero criticar, en cambio, es la manera en que Freud rechaza las críticas[15]. En efecto, estoy convencido de que Freud podía haber mejorado enormemente su teoría, si su actitud hacia las críticas hubiera sido diferente, sobre todo, su actitud hacia *las «críticas mal informadas»*, como les gusta llamarlas a los psicoanalistas. Y, no obstante, no cabe duda de que Freud era mucho menos dogmático que la mayoría de sus seguidores, que se sintieron inclinados a hacer una religión de la nueva teoría, dotada de todo lo necesario: mártires, herejes y cismas y que consideraron a todo crítico como un enemigo o al menos como «mal informado» (es decir, necesitado de ser analizado).

Esta actitud defensiva corre pareja con la actitud de buscar verificaciones, de encontrarlas en abundancia por todas partes, de rehusar admitir que ciertos casos no encajan en la teoría (y, al mismo tiempo, descartarlos por «no [ser] un problema de sueños sino de angustia»; es, de hecho, una «estratagema convencionalista» típica, como explico en la sección 20 de *L. I. C.*: una «inmunización», como la llama Hans Albert).

Una vez que se adopta esta actitud, *todo caso concebible se convertirá en un ejemplo verificador.* En 1919 ilustré esto con el siguiente ejemplo de dos casos radicalmente opuestos de conducta. Un hombre empuja a un niño al agua con la intención de ahogarle; y otro sacrifica su vida en un intento de salvar al niño. Cada uno de esos dos ejemplos de conducta puede explicarse fácilmente en términos freudianos y, por cierto, también en términos adlerianos. Según Freud, el primer hombre sufría de represión (por ejemplo de algún componente de su complejo de Edipo) mientras que el segundo había logrado sublimarlo.

---

ducción de Strachey. Véanse también las observaciones sobre «los profanos» en la décimo-cuarta de la *Introductory Lectures.*

[15] Véase, además de la nota 2 que acabo de citar, *op. cit.*, p. 135, también la nota añadida en 1925 a p. 160. Freud dice aquí de sus críticos que «no hacen mucho uso de su conciencia moral» (*Ges. Schriften*, III, p. 31), sugiriendo que les mueven sus «inclinaciones agresivas» cuando atribuyen la doctrina de que «todos los sueños tienen un contenido sexual» al «psicoanálisis». Sin embargo, ¿no pertenecía Otto Rank —que, como explica Freud, afirmaba precisamente eso— a las filas del «psicoanálisis»?

(Y, como escribió en una ocasión el psicoanalista S. Bernfeld, el psicoanálisis puede predecir que un hombre reprimirá o sublimará, pero no puede decir cuál de las dos cosas hará.) Según Adler, el primer hombre sufría de sentimientos de inferioridad (que le produjeron, quizá, la necesidad de demostrarse a sí mismo que se atrevía a cometer un crimen); y lo mismo le ocurría al segundo (cuya necesidad era demostrarse a sí mismo que se atrevía a arriesgar su vida). No puedo pensar en ningún ejemplo concebible de conducta humana que no pueda interpretarse en términos de cualquiera de las dos teorías y que no pueda ser reclamado, por cualquiera de ellas, como una «verificación».

[*(Añadido en 1980.) Ahora creo que la última oración del párrafo anterior es demasiado fuerte. Como me ha señalado Bartley, hay ciertos tipos de conducta que son incompatibles con la teoría de Freud, es decir, que la teoría de Freud los excluye. Así, la explicación que da Freud a la paranoia como un caso de homosexualidad reprimida parecería excluir la posibilidad de homosexualidad activa en un individuo paranoico. Pero esto no forma parte de la teoría básica que estaba criticando. Además, Freud podría decir de cualquier homosexual activo aparentemente paranoico que no es *realmente* paranoico o que no es *plenamente* activo.]

Que unos casos tan radicalmente opuestos hayan sido interpretados, de hecho, como verificaciones, puede mostrarse detalladamente analizando el tratamiento de Freud a ciertas objeciones a su teoría. En *La interpretación de los sueños,* Freud menciona «los sueños, muy frecuentes, que parecen contradecir mi teoría porque su tema es la frustración de un deseo o la ocurrencia de algo que claramente no se desea» (p. 157). Un grupo de esos «sueños contra-deseo», como él los llama, puede explicarse, dice, como sueños que cumplen el deseo del paciente de que la teoría de Freud sea equivocada. (Hay otro grupo que no nos preocupa aquí.) Así, la aparente falsación resulta ser una «verificación». Pero ¿qué ocurre con el caso radicalmente opuesto, de un paciente cuyos sueños se sueñan para complacer al analista y para confirmar su teoría y no para refutarla? Esos «sueños complacientes» (como los llama Freud en algunas ocasiones)[16] son también, desde luego, verificaciones; porque son satisfacción de deseos de la misma manera que los otros.

[16] Los «sueños complacientes» («*Gefälligkeitsträume*») o «sueños dóciles» —soñados para complacer o agradar al analista confirmando su teoría— los describe y discute Freud en *Ges. Schriften,* III, pp. 312-314 (correspondiendo a «Remarks on the Theory and Practice of Fream Interpretation», 1923, sección VII s., *Collected Papers* 5, 1950, pp. 141-145). Véase también la vigésimo-séptima y la vigésimo-octava de la *Introductory Lectures.*

Una actitud más crítica hacia esos «sueños complacientes» sería ésta: son, como dice el propio Freud, debidos a la sugestión del propio analista, al hecho de que el analista haya impuesto sus ideas a un paciente sugestionable. ¿No deberíamos considerar seriamente, entonces, la posibilidad de que otras «verificaciones clínicas» de las que les gusta hablar a los analistas, o incluso todas ellas, se deban a un mecanismo de este tipo? Y, la mera posibilidad de tal mecanismo, ¿no invalida tales «verificaciones»?

El propio Freud ve este problema y es interesante ver cómo lo trata[17].

«Quizá el analista se sobresalte al principio» —empieza Freud su discusión— «cuando se le recuerda por primera vez esta posibilidad», esto es, la de influenciar así a su paciente[18]. Es una observación interesante: «El analista», como Freud, «se sobresalta porque ve que todo su edificio de 'verificaciones clínicas' amenaza con desmoronarse». Pero la angustia del analista se calma tan pronto como se le dice que no es más que el «escéptico» el que le recuerda esta chocante posibilidad: «El escéptico puede decir que esas cosas aparecen en el sueño porque el que sueña sabe que tiene que producirlas, que el analista las espera», escribe Freud; y añade: «el propio analista pensará, con justicia, de modo diferente»[19].

Sin duda lo hará. Pero ¿por qué «con justicia»? Freud no da ninguna razón. Por el contrario, cuando el escéptico reaparece tres páginas más adelante por última vez —entonces le llama «alguien»— ni siquiera el propio Freud «piensa de un modo diferente»; porque ahora escribe: «Si alguien mantuviera que la mayoría de los sueños que pueden utilizarse en un análisis son, de hecho, sueños complacientes que se han producido debido a la sugestión [del analista], *entonces, nada puede decirse contra esta opinión desde el punto de vista de la teoría analítica.* En este caso sólo necesito referirme a las consideraciones en mis conferencias introductorias [*Introductory Lectures*], donde... muestro lo poco perjudicada que resulta la credibilidad de nuestros resultados por una comprensión del efecto de la sugestión, en nuestro sentido»[20]. Me temo que la referencia a *Las conferen-*

[17] *Op. cit.*, *Ges. Schriften*, III, pp. 310-314.
[18] *Op. cit.*, p. 130.
[19] *Op. cit.*, p. 311. (*Coll. Papers* 5, p. 142.)
[20] *Op. cit.*, p. 314. (Los subrayados son míos. *Cf. Coll. Papers*, p. 145.) La referencia es al punto cuatro de la décimo-quinta de las *Introductory Lectures* que proporciona, como mucho, un argumento circular como réplica y a la última (vigésimo-octava) que proporciona un argumento que sólo muestra que algunos de los principios más generales del análisis están apoyados por evidencia independiente, de modo que no todo puede deberse a la sugestión.

cias introductorias difícilmente ayudará a nadie a superar la contradicción entre las dos últimas citas. Si alguien puede pensar con sentido crítico, debe permanecer en estado de «sobresalto»; sobre todo si lee, entre las líneas de la décimo-quinta de esas *Conferencias introductorias* (*cf.* las seis primeras líneas del punto 4), que el sobresalto tuvo su origen en el descubrimiento de que los pacientes de Freud, Adler y Stekel soñaban, respectivamente, «principalmente con impulsos sexuales,... de dominio... [y] de re-nacimiento», adaptando de este modo como lo expresa Freud «el contenido de sus sueños a las teorías favoritas de sus médicos».

Pero volviendo de *Las conferencias introductorias* al pasaje que me indujo a citarlas, el único argumento digno de ese nombre en esas cuatro páginas de disculpas es *el argumento del rompecabezas de piezas irregulares* (pp. 312-313 y 143, respectivamente). Afirma que, si el analista logra recomponer la totalidad de la intrincada imagen «de modo que el dibujo adquiera sentido, y no queden huecos en ningún lado... entonces sabe que ha hallado la solución y que no hay otra».

Nada podría ser más peligroso que este argumento si, como en el contexto actual, se usa para disipar las dudas del analista respecto a los resultados de la sugestión. Porque de lo que tenía miedo el analista, para empezar, era precisamente de la posibilidad de que el rompecabezas pudiera componerse bajo presión, estrujando las piececitas (que resultan ser elásticas o plásticas y no rígidas) para encajarlas en su sitio, o quizá mediante la sugestión inconsciente a un paciente dócil para que produzca algunas piezas nuevas, hechas especialmente a la medida, para que encajen bien en los diversos «huecos».

Incluso sin esta objeción decisiva, el argumento del rompecabezas sólo es aceptable si tenemos ante nosotros una teoría que pueda ser contrastada rigurosamente: las demás teorías *siempre* pueden hacer que sus rompecabezas encajen. Consideremos, por ejemplo, las interpretaciones de la historia en términos de luchas de razas o de luchas de clases: qué bien «resuelven» todas ellas el rompecabezas de la historia y de la política actual. Lo mismo puede decirse de la interpretación astrológica de la historia o de la interpretación que hace Homero de ella, en términos de las trifulcas domésticas en el Monte Olimpo, o de la interpretación del Antiguo Testamento en términos de

(Estoy dispuesto a conceder esto; lo que es quizá la evidencia más sorprendente de este tipo puede encontrarse en la *República* de Platón, por ejemplo, pp. 571-575. Estos pasajes no los menciona Freud. Estos y otros se discuten en *La sociedad abierta* [*O. S.*], nota 59 al capítulo 10.) El pasaje en la conferencia vigésimo-octava también hace uso de lo que es esencialmente el argumento del rompecabezas de piezas irregulares; véase *infra*.

culpa colectiva, castigo y expiación. Cada una de ellas logra «resolver» su rompecabezas. Pero su creencia —y la de Freud— «de que no hay otra solución» carece de fundamento: todas ellas tienen éxito. (Y también lo tuvieron Adler y Stekel.)

No quiero ser mal interpretado. Pienso que *La interpretación de los sueños* de Freud es un gran logro. Con todo, tiene un carácter más parecido al del atomismo anterior a Demócrito —o quizá al de la colección de relatos de Homero sobre el Olimpo— que al de una ciencia contrastable. Muestra ciertamente que incluso una teoría metafísica es infinitamente mejor que la carencia de teorías; y, supongo, es un programa para una ciencia psicológica, comparable al atomismo y al materialismo, o a la teoría electromagnética de la materia, o a la teoría de campos de Faraday, que fueron, todas ellas, programas para la ciencia física. *Pero es un error fundamental creer que, porque está «verificándose» constantemente, tiene que ser una ciencia basada en la experiencia.*

Siempre acompaña al verificacionismo un dogmatismo peligroso. Yo no creo que la pregunta «¿cuál es la naturaleza esencial de los sueños?» sea una buena pregunta; pero si se plantea, entonces, las respuestas diferentes de la teoría de Freud sobre la satisfacción de los deseos parecen ser por lo menos tan apropiadas como ésta. Por ejemplo, todo el material de Freud y sus análisis podrían encajar muy bien en la siguiente réplica: «todos los sueños son consecuencia de conflictos: de deseos en conflicto o de conflictos entre los deseos y los obstáculos que amenzan con frustrarlos y que crean preocupaciones y problemas»[21]. Ahora bien, como en un sueño los deseos no pueden expresarse más que en forma de una representación de lo que se desea[22] —es decir, de su satisfacción— una representación de esta satisfacción se encontrará en la mayoría de los sueños. Sin embargo, aunque algunos sueños pueden culminar en una satisfacción, el conflicto y la frustración están siempre representados con fuerza (incluso en los sencillos sueños de la niñez y en los sueños de hambre); y se convierten en dominantes en un sueño de angustia, que *no* tiene por qué ser un síntoma de neurosis de angustia[23].

---

[21] Los sueños que representan trabajo sobre un problema se discuten en *The Interpretation of Dreams*, nota 2 (añadida en 1919), en p. 181, donde Freud se refiere a experimentos realizados por Pötzl (véase también p. 569). Samuel Butler había previsto algunos de los resultados de Pötzl en un espléndido pasaje en *Erewhon Revisited* (1901), capítulos 27 y 28. «Me gustaría que alguien escribiese un libro sobre sueños», escribió Butler en el capítulo 28, ignorando que Freud había publicado su libro dos años antes.

[22] Esta idea está tomada de la última oración de la sección III de Freud, *On Dreams*, 1901; *Standard Edition* 5, pp. 629 ss.

[23] Véase Butler, *loc. cit.*, para una interesante descripción y análisis de un

Nada más lejos de mi intención que ofrecer esta teoría —que, en cualquier caso, se lo debería todo a Freud— como alternativa a la suya[24]. Lo que quiero señalar es que Freud no discute en ningún lugar una teoría alternativa (tal como la esbozada aquí) que tome nota del simple hecho, ahora admitido, de que *los sueños de angustia constituyen una refutación de la fórmula general de satisfacción de deseos,* como sugirieron hace mucho tiempo los lectores «obstinados» y los críticos «mal informados». En ningún lugar compara Freud su teoría con una rival prometedora, sopesando la una contra la otra a la luz de la evidencia; y nunca la critica: tiene su teoría y trata de verificarla; la hace encajar, mientras sea posible y —como lo muestra el ejemplo del sueño de angustia— incluso más allá de lo que él mismo pensó posible cuando publicó por primera vez su gran libro, *La interpretación de los sueños.*

Estas fueron las razones, más o menos, que me llevaron, en 1919, a rechazar las pretensiones de freudianos, adlerianos y marxistas respecto a que sus teorías estaban «basadas en la experiencia» del mismo modo que las teorías de las otras ciencias, la neurología experimental o, digamos, la bioquímica. Rechacé sus pretensiones porque encontré que sus teorías no satisfacían el criterio de contrastabilidad o refutabilidad o falsabilidad. Hoy ese criterio está siendo ampliamente aceptado como criterio de demarcación; pero las tres teorías mencionadas rara vez se discuten en relación con él. En cambio, siguen discutiéndose en términos de evidencia confirmadora, de «verificaciones».

---

sueño de angustia. Unos sueños de hambre —de Shackleton y Wilson— que no encajan en la teoría de Freud, pero que encajan en la propuesta aquí, son los referidos por el capitán Scott en *The Voyage of Discovery* (anotación en el diario con fecha 22 de diciembre de 1902): «Mis compañeros tienen muy malos 'sueños de comida'; han llegado a convertirse en la conversación normal durante el desayuno. Parece ser una especie de pesadilla; o bien están sentados ante una mesa bien servida y tienen los brazos atados, o cogen un plato y se les resbala de la mano, o están en el acto de llevarse a la boca un pedazo exquisito cuando caen por un precipicio. Cualesquiera que sean los detalles, algo se interpone en el último momento y se despiertan».

[24] Como otra alternativa a la fórmula de Freud de satisfacción de deseos, J. O. Wisdom propuso, en el artículo al que se refiere la nota 5 *supra*, que se reelaborase y se reinterpretase. Atribuyendo mucho peso a la idea de Freud de deseos de castigo (una idea de la que el propio Freud se había servido muy poco), Wisdom porpone explicar *todos* los sueños —incluso los de angustia— en términos de satisfacción de deseos (el propio Wisdom prefiere el término «satisfacción de necesidades», pero no veo ninguna razón, desde un punto de vista psicoanalítico, para que no pueda haber un «deseo inconsciente» que correspondan a cada «necesidad» en el sentido de Wisdom, incluido lo que él llama «necesidades de castigo»). Puede decirse que la teoría de Wisdom admite conflictos en la explicación de los sueños, pero, por lo que yo veo, sólo admite *un* tipo de conflicto: el debido a sentimientos de culpabilidad.

Así es como vi por primera vez el problema de demarcación. En el contexto actual, no importa si tengo o no razón respecto a la refutabilidad de ninguna de esas tres teorías: aquí sólo sirven como meros ejemplos, ilustraciones. Porque mi objetivo es mostrar que mi «problema de demarcación» fue desde el comienzo el problema práctico de evaluar las teorías y de juzgar sus afirmaciones. No era, por cierto, un problema de clasificarlas o de distinguir ciertas disciplinas llamadas «ciencia» y «metafísica». Era, más bien, un problema práctico urgente: ¿bajo qué condiciones es posible un *recurso crítico a la experiencia,* uno que dé fruto?

19. CONTRASTABILIDAD, PERO NO SENTIDO

He contado parte de la historia de cómo llegué a formular por primera vez el problema de demarcación porque quiero mostrar que no fueron ni el dogma filosófico de la falsabilidad ni las dificultades filosóficas de la verificabilidad los que me llevaron a él. Más bien fue un problema urgente: el de decidir si una teoría era aceptable: si era argüible por medio de argumentos empíricos (es decir, argumentos que apelasen a la observación y al experimento) y si esos argumentos debían considerarse como contrastaciones serias. Mi problema resultó ser un problema lógico, un problema metodológico e incluso un problema de la propia ciencia, todo a la vez. Porque es tarea del científico juzgar las teorías; y una forma de emitir juicio sobre una teoría es decir que no puede ser juzgada con los patrones científicos ordinarios (es decir, valorando cómo resiste a las contrastaciones) porque es irrefutable y, por tanto, incontrastable.

De ahí que yo sugiriese que la contrastabilidad o refutabilidad o falsabilidad se aceptase como criterio para juzgar el carácter científico de los sistemas teóricos; es decir, *como un criterio de demarcación entre la ciencia empírica por un lado y la matemática pura, la lógica, la metafísica y la pseudo-ciencia por otro.*

Nunca se me ocurrió, ni en aquellos años ni posteriormente, proponer la contrastabilidad o refutación o falsabilidad como un criterio de *sentido* (en contraposición a «absurdo sin sentido»); y cuando oí por primera vez, en 1927 más o menos, que el Círculo de Viena había aceptado la verificabilidad como un criterio de sentido[25], objeté al instante sobre ese procedimiento,

[25] La formulación se debe a F. Waismann (véase la nota 2 a la sección 6 de *L. I. C.*), aunque puede decirse que la idea es, más o menos, de Wittgenstein, supongo. Con frecuencia se olvida ahora que una expresión se llama

basándome en dos argumentos enteramente diferentes: en primer lugar, porque tomar el *sentido* como criterio de demarcación equivalía a poner a la metafísica la etiqueta de galimatías sin sentido, un dogma que yo me sentía incapaz de aceptar; y, en segundo lugar, porque la *verificabilidad* se proponía como criterio de significado o sentido o significación y, por tanto, indirectamente, como criterio de demarcación: una solución que era totalmente inadecuada y, de hecho, lo contrario de lo que hacía falta. Porque yo podía mostrar que era un criterio *demasiado estrecho y demasiado amplio* al mismo tiempo: (inintencionadamente) declaraba carentes de sentido las teorías científicas y con ello las situaba (también inintencionadamente) en el mismo plano que la metafísica (*cf.* sección 4 de *L. I. C.* y Apéndice *II).

Además, era «verificacionista» en el importante sentido de que pasaba por alto el hecho de que la discusión científica (como un cierto tipo de discusión racional) era discusión *crítica* y que su actitud fundamental era buscar refutaciones más que buscar verificaciones o confirmaciones.

La ancha línea de demarcación entre ciencia empírica por un lado y pseudo-ciencia o metafísica o lógica o matemática pura por otro, tiene que trazarse a través del corazón mismo de la región del sentido —con teorías con sentido a ambos lados de la línea divisoria— y no entre regiones de sentido y sinsentido. Rechazo, sobre todo, el dogma de que la metafísica tiene que ser absurda. Porque como hemos visto, algunas teorías, como el atomismo, fueron durante mucho tiempo incontrastables e irrefutables (e, incidentalmente, no verificables, también) y, mientras tanto, «metafísica». Pero más tarde pasaron a formar parte de la ciencia[26]. Y otras sufrieron el destino

---

«carente de sentido» o «absurda» en el sentido, en cierto modo técnico (debido a Russell y aceptado por Wittgenstein y el Círculo de Viena), que estamos discutiendo aquí, sólo si es un puro galimatías. Un buen ejemplo de una pseudocláusula carente de sentido del *Tractatus* de Wittgenstein es: «Sócrates es idéntico». Contrariamente al sentido habitual de «absurdo», la afirmaciones claramente tontas como «Tres tristes triunviros trogloditas en tropa trepan tras el trapecio trigonométricamente trastocados en trípodes triclínicos», o «3 + 11 = 33», no son «absurdas» o «carentes de sentido» en este sentido técnico, sino *falsas*. Además, *toda* proposición falsa, por tonta que sea, siempre tiene significado (es decir, sentido); y su negación, por trivial que sea, es siempre verdadera. (Respecto a la clasificación de Russell de las expresiones en verdaderas, falsas y absurdas, y su influencia sobre Wittgenstein, véanse también mis artículos «La naturaleza de los problemas filosóficos y sus raíces en la ciencia» y «Autorreferencia y significado en el lenguaje común», ambos en *El desarrollo del conocimiento científico* [*C. & R.*].)

[26] El carácter metafísico de la «teoría corpuscular» de la materia (el atomismo) se explica brevemente en la sección 23, *infra* (en el texto que sigue a la nota 37).

contrario. Es claramente inadecuado describirlas como insensateces. Estoy dispuesto a admitir que algunos metafísicos (estoy pensando, en especial, en Hegel y los hegelianos) se han complacido en decir insentateces y, lo que es peor, insensateces pretenciosas. Sin embargo, los científicos no están del todo libres de esa enfermedad. En cualquier caso parece preferible no tomarse demasiado en serio el sin-sentido. Desde luego, es algo insano y un poco necio escoger el problema de lo insensato, o de la carencia de sentido, como problema básico de la propia filosofía y su exposición como la principal o, quizá, la única tarea de uno: el análisis de significados, como el psicoanálisis, puede convertirse fácilmente en «un padecimiento que se confunde a sí mismo con su cura»[27].

No podría yo haber sido más explícito en mi rechazo de todo el problema del sentido o significatividad *versus* el sentido o la carencia de significado, de lo que fui desde el principio, como puede ver cualquier lector de mi *L. I. C.* Lo acusé de pseudo-problema, como tentativa equivocada de formular el problema de demarcación y como solución equivocada a ese problema. Y he reafirmado consecuente y repetidamente mi posición[28]. Sin embargo, a pesar de todo esto, se me llama «el positivista que propuso modificar el criterio de *sentido* según la verificabilidad y sustituirlo por un criterio de *sentido* que dependa de la falsabilidad». En 1949, a esta etiqueta se le había añadido un aditamento: «Pero entiendo que *ahora* no se llamaría a sí mismo positivista»[29]. Sólo puedo decir que siempre he criticado severamente al positivismo y que nunca he cambiado de opinión en estos asuntos; y pido al lector que se cerciore de este hecho leyendo el Apéndice *I (1933) y las secciones 4, 10, 79 y 85 de mi *L. I. C.* (1934).

Estoy resignado al hecho de que, a pesar de todo, llevaré esta etiqueta hasta el final de mis días. (Y quizá hasta después,

---

[27] Me refiero a un aforismo del poeta vienés Karl Kraus que escribió «*Die Psycoanalyse ist jene Geisteskrankheit die sich für ihre Therapie hält*» (*Nachts*, 1924, p. 80). O, traducido libremente: «El psicoanálisis es el padecimiento que se confunde a sí mismo con su cura».

[28] Véase, aparte del Apéndice *I, y las secciones 4 y 10 de mi *L. I. C.*, mi *La sociedad abierta y sus enemigos* [*O. S.* (1945 y ediciones posteriores)], sobre todo las notas 46, 51 y 52 al capítulo 11, y otras referencias que allí se hacen; también la nota 2 al capítulo I; véanse también mis artículos «Indeterminism in Quantum Physics and in Classical Physics», *The British Journal for the Philosophy of Science* 1, 1950-51; «The Nature o Philosophical Problems, etc.», *ibid.* 3, 1952; «Tres concepciones sobre el conocimiento humano», 1956; y «Phylosophy of Science: A Personal Report», 1957. [*Cf. Búsqueda sin término (U. Q.)*, secciones 8 y 16, y *El desarrollo del conocimiento científico (C. & R.)*, capítulos 1, 2, 3. *Ed.*]

[29] *Cf.* nota 3 en p. 76 de W. Kneale, *Probability and Induction*. [*Cf. Búsqueda sin término (U. Q.)*, sección 17. *Ed.*]

si la historia de la filosofía continúa prestándonos atención a los últimos rezagados de la ilustración; lo que parece ser una posibilidad remota, en vista de la ya abrumadora y aún creciente demanda de un mesianismo filosófico irracional y antirracional *à la* Heidegger, por un lado, y de un método filosófico «matemáticamente exacto» por el otro)[30]. Sin embargo, me propongo repetir, hasta el fin de mis días, que la confusión del problema de demarcación con el de sentido es uno de los mayores errores de la escuela de pensamiento positivista.

20. ENUNCIADOS NO CONTRASTABLES

La suposición errónea, corregida en la sección precedente, de que mi criterio de demarcación trata de ser un criterio de sentido, puede dar lugar a mucha confusión. No sólo malinterpreta mi intención, sino que es formalmente incompatible con mi teoría. Así, si se superpone a mi teoría, el resultado será auto-contradictorio. Esto puede mostrarse como sigue.

Si tomamos cualquier enunciado con sentido y formamos su *negación,* el resultado será, obviamente, otro enunciado con sentido. La negación puede formarse anteponiendo al enunciado las palabras: «No ocurre que». Tomemos ahora una expresión sin sentido y antepongamos esas palabras. El resultado será, obviamente, otra expresión sin sentido. Llegamos así al siguiente requisito, que tiene que se satisfecho por cualquier criterio coherente y adecuado de significado y de carencia de significado: *La negación de un enunciado con sentido tiene que tener sentido y la de una expresión sin sentido, de una sucesión de palabras sin sentido, tiene que carecer de sentido.*

Este requisito resultará violado si se interpreta mi criterio de demarcación como un criterio respecto a la significatividad de una expresión.

Porque la negación de un enunciado universal falsable es un enunciado existencial infalsable. Por ejemplo: «todos los cuervos son negros» es falsable, puesto que podemos encontrar un día un cuervo blanco. Pero su negación es el enunciado existencial, «existe (en algún lugar del universo, en el pasado, presente o futuro) un cuervo que no es negro»; y este enunciado existencial no es contrastable, puesto que no puede ser falsado por ninguna cantidad de informes sobre observaciones. Así pues, si mi criterio de demarcación fuese un criterio de significado, este enunciado existencial tendría que ser considerado sin sentido, aunque es la negación de un enunciado universal con sentido.

[30] *Esta »exactitud«* es decepcionante. Véanse las Observaciones Preliminares al capítulo III.

En consecuencia, el requisito resulta violado. (En efecto, el criterio de verificabilidad lo violaba, puesto que era un criterio de sentido.) Esto no muestra que mi criterio de demarcación sea defectuoso. Sólo muestra que de cualquier intento de interpretar mi criterio de demarcación entre la ciencia y la metafísica como un criterio de sentido, se derivan consecuencias absurdas. No es extraño que algunos filósofos —los que no eran capaces constitucionalmente de concebir una demarcación entre la ciencia y la metafísica que no fuera al mismo tiempo una demarcación entre el sentido y el sin sentido— encontrasen insatisfactorio mi criterio de demarcación.

Ahora bien, el status de los enunciados existenciales en mi teoría es bastante simple y claro. Respecto a su significado, no caben dudas: tienen tanto sentido como sus negaciones, los enunciados universales. Así pues, el requisito formulado más arriba se satisface de manera normal (aunque estoy lejos de tratar de crear una teoría del significado).

Respecto a la cuestión del carácter científico de los enunciados existenciales, es importante recordar que mi criterio es aplicable a *sistemas teóricos* más que a enunciados sacados del contexto de un sistema teórico (cómo lo señalé desde el principio; véase *L. I. C.*, Apéndice I). Esto se debe al hecho de que me interesa mucho más *lo que* dice una teoría que *cómo* lo dice. Pero una y la misma teoría puede formularse de muchas maneras diferentes, con la ayuda de un número mayor o menor de hipótesis (premisas). Por tanto, afirmé en *L. I. C.* (sección 15) que, según mi criterio, algunos enunciados existenciales son científicos, es decir, los que pertenecen a un contexto contrastable. Un ejemplo que doy allí es «existe un elemento con el número atómico 72». Es científico como parte de una teoría sumamente contrastable, y una teoría *que da indicaciones de cómo encontrar ese elemento.* Si, por otro lado, tomásemos aisladamente este enunciado, o como parte de una teoría que no nos diese ningun indicio de cómo y dónde encontrar ese elemento, entonces tendríamos que describirlo como metafísico, simplemente porque no sería contrastable: ni siquiera un elevado número de fracasos en la detección del elemento podría considerarse como incapacidad del enunciado de pasar una contrastación, porque nunca podríamos saber si el próximo intento podría producir el elemento y verificar concluyentemente la teoría.

Por tanto, algunos enunciados existenciales serán contrastables y otros incontrastables. Y a veces esto dependerá del contexto y quizá cambie al cambiar el contexto[31].

[31] [*Cf.* «Replies to my Critics», en P. A. Schilpp, *The Philosophy of Karl Popper*, Vol. II, sección II, pp. 976-1013, y pp. 1037-9. *Ed.*]

21. EL PROBLEMA DE «ELIMINAR» LA METAFISICA

El hecho de que mi criterio de demarcación entre la ciencia y la metafísica esté concebido para ser aplicado sólo a *sistemas teóricos* es un punto de considerable importancia para su evaluación, más importante que la cuestión de los enunciados existenciales.

Este punto, que destaqué desde el principio, distinguía mi criterio del de los positivistas; porque ellos creían que su criterio de significado era aplicable a cualquier *expresión lingüística* sin referencia a su *contexto*. (Ellos pensaban que lo único que se necesitaba era un conocimiento de las reglas de la lengua a la que pertenecía la expresión.) Creían que su criterio de significado les permitiría *detectar el absurdo* donde quiera que apareciera. Así, algunas veces describían como objetivo suyo la «eliminación de la metafísica por medio del análisis del lenguaje»[32]. Y creían que tenían un método, una técnica, que les permitiría eliminar los elementos metafísicos —es decir, el absurdo— también de las teorías científicas.

Yo no creo que la metafísica sea absurda, y no creo que sea posible eliminar todos los «elementos metafísicos» de la ciencia: están demasiado estrechamente entretejidos con el resto. No obstante, creo que siempre que sea posible encontrar un elemento metafísico en la ciencia que *pueda* ser eliminado, la eliminación será para bien. Porque la eliminación de un elemento no contrastable de la ciencia suprime un modo de evitar las contrastaciones; y eso tenderá a aumentar la contrastabilidad o refutabilidad del resto de la teoría. Y, en un número de casos, una teoría científica ha ganado considerablemente con el descubrimiento de que había en ella elementos metafísicos que podrían ser eliminados, y por el intento de eliminarlos.

Mi criterio de demarcación, no obstante, no se ha cocebido como instrumento para la detección de dichos elementos. No intento dar a entender que no puede usarse para ello. En realidad, lo he encontrado bastante útil en unos cuantos casos. Pero no creo que la detección de elementos metafísicos y su eliminación de la ciencia pueda convertirse en parte de una rutina o de una *técnica*.

Los positivistas adoptan una perspectiva diferente respecto a la posibilidad de desarrollar una técnica así. Para un positivista, ello supondría simplemente buscar errores gramaticales y otros errores lingüísticos similares (que, después de todo, cometen la mayor parte de los autores).

[32] *Este era el programa de investigación de Carnap.*

Pero donde el contenido de las teorías adquiere toda su importancia, como cuando se mejora una teoría eliminando elementos metafísicos *con sentido*, la tarea formará parte de la crítica racional; y la crítica racional es siempre un proceso imaginativo y creativo y no una mera técnica. La «eliminación de elementos metafísicos» nunca consiste en la mera omisión de una o dos oraciones, sino que siempre supone una reconstrucción de la teoría, inspirada, por regla general, por una nueva idea respecto a su interpretación.

Deberíamos recordar también que una y la misma teoría pueden interpretarse o bien en el sentido convencionalista (como un conjunto de definiciones irrefutables), o empíricamente, como señalé en mi discusión del convencionalismo (*L. I. C.*, secciones 19 ss.). Esto muestra cuánto depende de la *interpretación* de una teoría, si queremos juzgar su carácter empírico aplicando nuestro criterio de demarcación; y muestra, en consecuencia, que la tarea de descubrir «elementos metafísicos» no puede resolverse sólo escudriñando el formalismo. De hecho, es sólo parte de la tarea más general de perfeccionar una teoría, por medio de la crítica, y por el método de prueba y error.

Un caso interesante, en relación con esto, es la crítica que Berkeley hizo de Newton. Berkeley logró no sólo presentar una excelente crítica del carácter «oculto» o metafísico del espacio absoluto de Newton, sino que incluso anticipó, en su *De motu*, la famosa sugerencia de Mach de que se reformase la teoría, recogida más tarde por Einstein en su teoría de la relatividad general[33]. Pero deberíamos recordar que en conjunción con la teoría ondulatoria de la luz, la teoría del espacio y el tiempo absolutos se convirtió en contrastable, de modo que el experimento de Mechelson pudo ser interpretado como una refutación. No fue, por tanto, el carácter intrínsecamente metafísico de esos conceptos lo que hizo deseable su eliminación, sino el hecho de que, dado *sólo* el contesto de la mecánica de Newton, representaban elementos incontrastables.

[*Tal vez pueda mencionarse que mi interpretación de las fórmulas de la indeterminación de Heisenberg como relaciones de dispersión fue al mismo tiempo un intento de criticar el positivismo machiano de Heisenberg («observables») y de eliminar lo que yo consideraba como su dogmatismo metafísico: su teoría de que las fórmulas de indeterminación señalaban los límites del conocimiento humano.]

[33] Véase mi artículo «Nota sobre Berkeley como precursor de Mach», *British Journal for the Philosophy of Science* 4, 1953, pp. 26 ss., en mi *El desarrollo del conocimiento científico (C. & R.)*.

22. LA ASIMETRIA ENTRE FALSACION Y VERIFICACION

Como señalé en *L. I. C.*, hay una asimetría fundamental entre la falsación empírica y la verificación empírica. Aunque algunos de mis críticos han negado la existencia de esa asimetría, sus argumentos fueron previstos y plenamente discutidos en *L. I. C.*

Esta fundamental asimetría no puede negarse seriamente, pienso: un conjunto de enunciados singulares de observación («enunciados básicos», como los llamé) pueden falsar o refutar a veces una ley universal; pero no puede de ningún modo verificar una ley, en el sentido de establecerla. Precisamente el mismo hecho puede expresarse diciendo que un conjunto de enunciados singulares de observación puede verificar un enunciado existencial (lo que equivale a falsar una ley universal) pero no que puede falsarlo. Esta es la situación lógica fundamental; y muestra una llamativa asimetría.

De las diversas objeciones que se han planteado contra mi afirmación de que existe esta asimetría —y con ellas, contra mi criterio de demarcación— la que llama más la atención a primera vista es la siguiente: Cuando falsamos un enunciado, automáticamente verificamos su negación, porque la falsación de un enunciado *a* puede interpretarse siempre como la verificación de su negación no-*a*. En consecuencia, siempre podemos hablar, si queremos, de verificación en lugar de hablar de falsación, y viceversa: la diferencia entre esas dos formas de expresar las cosas es meramente verbal y, por tanto, son, por razones lógicas, completamente simétricas.

Por ejemplo, si podemos describir una *contrastación empírica* como intento de falsación o como una búsqueda de un ejemplo positivo (del enunciado *a*), entonces también podemos describirlo como un intento de verificación o como una búsqueda de un ejemplo positivo (del enunciado no-*a*). Del mismo modo, cualquier obstáculo a la verificación de un enunciado *a*, debe, por razones lógicas, constituir un obstáculo para la falsación de no-*a*, y viceversa. Así, se dice que es difícil verificar un enunciado existencial tal como «existe una máquina en movimiento perpetuo», es decir, «una máquina que continúa emitiendo energía sin absorber nunca energía de su entorno», puesto que tendríamos que registrar todo el mundo en su busca (y, de paso, examinar toda candidata durante un período indefinido de tiempo); pero, obviamente, debe ser igualmente difícil falsar la negación de este enunciado existencial, es decir, el enunciado universal «todas las máquinas que continúan emi-

tiendo energía deben, después de un tiempo finito, absorber energía de su entorno». Porque la verificación de un enunciado no es sino la falsación del otro.

Ahora bien, de estas premisas obvias se deduce, concluye mi crítico, que por razones puramente lógicas, es inútil diferenciar entre falsación y verificación o deslindar una clase de enunciados «falsables» o «contrastables» como «científicos» y distinguirlos de otra clase de enunciados no falsables (aunque quizá verificables unilateralmente), a los que llamamos «metafísicos». Así, el falsacionista que afirma que el enunciado «todos los cisnes son blancos» puede falsarse, tendrá que admitir que cada falsación o refutación de este enunciado universal será equivalente a la verificación y aceptación del enunciado existencial «existe un cisne no blanco». Debe ser un error, por tanto, llamar «científico» al enunciado universal y «metafísico» al enunciado existencial. (La lección que suele sacarse de esto es que la distinción entre enunciados científicos y enunciados metafísicos no depende de cosas como la contrastabilidad —es decir, la relación entre enunciados— sino más bien de los *conceptos* —observables o no— que ocurren en los enunciados)[34].

En respuesta a esta crítica, quiero decir que acepto todas las premisas de mi crítico como verdaderas, de hecho, como trivialmente verdaderas; pero rechazo todas sus conclusiones (que están enunciadas en el párrafo anterior) ninguna de las cuales, por cierto, se desprende de sus premisas.

En primer lugar me gustaría quitar de en medio un punto, un punto con el que estoy de acuerdo, puesto que pertenece a las premisas de mi crítico, pero que revela, no obstante, un malentendido. Es la referencia, en la parte final de las premisas de mi crítico, a «obstáculos» o «dificultades» que a veces impiden la verificación de un enunciado puramente existencial tal como «existe una máquina en movimiento perpetuo». Naturalmente, lo que dice mi crítico es verdad: la verificación empírica de ese enunciado es, por razones obvias, exactamente igual de difícil o exactamente igual de fácil que la falsación empírica de la ley universal que es su negación. Pero yo nunca me he preocupado por esta «dificultad», y nunca me he referido a ella o sacado ninguna conclusión de ella. Yo no llamo «metafísico» a un enunciado puramente existencial aislado porque sea «difícil» de verificar, sino porque es *lógicamente imposible falsarlo* empíricamente o contrastarlo. Y, naturalmente, siempre he hecho hincapié en que la imposibilidad lógica de falsar un enunciado existencial de este tipo es exactamente lo mismo que la imposi-

[34] Véanse especialmente las secciones 24 y 25; véase también la nota 93 a la sección 11.

bilidad lógica de verificar su negación universal. La referencia de mi crítico a obstáculos y dificultades, por tanto, no hace al caso. Además, parece revelar una actitud verificacionista: los verificacionistas, al parecer, no pueden imaginar ninguna dificultad en relación con enunciados puramente existenciales, aparte de la dificultad de verificarlos.

El punto relativo a los «obstáculos» o «dificultades» puede descartarse, pues, como no pertinente; y puedo proceder a ocuparme de otros puntos, quizá más pertinentes.

Sin duda, puede decirse que los problemas de verificación y falsación son, en ciertos aspectos, «simétricos». El hecho de que haya ciertas simetrías aquí no impide la existencia de una asimetría fundamental, del mismo modo que la existencia de una simetría de largo alcance entre números positivos y negativos no impide una asimetría fundamental en el sistema de los números enteros: por ejemplo, que un número positivo tiene raíces cuadradas reales, mientras que un número negativo no la tiene.

Así pues, uno puede ciertamente decir que la falsabilidad y la verificabilidad son «simétricas» en el sentido de que la negación de un enunciado falsable debe ser verificable y viceversa. Pero este hecho, que resalté repetidas veces en *L. I. C.* (donde incluso describí los enunciados universales como enunciados existenciales negativos), y que es todo lo que mi crítico establece en sus premisas, no es argumento en contra de la fundamental asimetría cuya existencia he señalado.

Esta asimetría tiene un aspecto puramente lógico y también uno metodológico o heurístico.

En cuanto a su aspecto lógico, no hay duda de que un enunciado universal (unilateralmente falsable) es lógicamente mucho más fuerte que el correspondiente enunciado existencial (unilateralmente verificable). Porque lo que sigue es una regla lógica bien conocida: de un enunciado universal, relativo a todas las cosas de un cierto tipo, o a todos los elementos de un cierto universo no vacío de discurso,

1) todas las cosas tienen la propiedad *P*,

podemos deducir, que para cualquier cosa individual *a* que pertenezca a este tipo o universo, que

2) la cosa *a* tiene la propiedad *P*,

y de 2), a su vez, podemos deducir que

3) existe una cosa que tiene la propiedad *P*.

De modo que 1) entraña 2) y 3), y 2) entraña 3). Pero 3) no entraña ni 1) ni 2), y 2) no entraña 1).

O, dicho de otro modo, 1) es lógicamente más fuerte que 2) y 3) y 2) es lógicamente más fuerte que 3).

Esta es la fuente de la importante asimetría en el caso de los enunciados universales unilateralmente falsables y los enunciados existenciales unilateralmente verificables; y la situación es la misma con enunciados más complejos. (Véase sección 24.)

Debido a su poder lógico, los enunciados universales pueden ser importantes como *hipótesis explicativas*: pueden explicar (especialmente en conjunción con condiciones iniciales singulares) sucesos o enunciados singulares. Los enunciados puramente existenciales, por otro lado, aislados, o incluso en conjunción con enunciados singulares, son normalmente demasiado débiles para explicar nada.

Esta es la razón de que los científicos estén interesados en hipótesis universales y no en hipótesis existenciales (aisladas).

Esto nos lleva al aspecto metodológico o heurístico de la asimetría: a la diferencia entre la actitud crítica o falsacionista y la actitud verificacionista.

La concepción del verificacionismo sobre la ciencia es algo parecido a esto: idealmente, la ciencia consiste en todos los enunciados verdaderos. Como no los conocemos todos, debe consistir al menos en todos aquellos que hemos verificado (o quizá «confirmado» o mostrado que son «probables»). Así pues, los enunciados existenciales verificados deben, *por esta razón*, pertenecer a la ciencia.

La actitud del falsacionista es diferente. Para él, la ciencia consiste en osadas hipótesis explicativas, «osadas» en el sentido de que afirman tanto que fácilmente pueden resultar falsas. Y hace todo lo que puede para encontrarles faltas, con la esperanza de detectar y eliminar a las candidatas defectuosas que aspiran a la posición de teoría explicativa y también con la esperanza de ganar de ese modo mayor penetración. Respecto a los enunciados puramente existenciales, no le interesan a causa de su debilidad y porque no pueden ser falsados a menos que formen parte integral de un sistema teórico. El falsacionista está dispuesto a admitirlos en la ciencia *si* son entrañados por un enunciado básico aceptado; pero aun así, su interés estriba solamente en el hecho de que su aceptación es equivalente al rechazo de sus negaciones universales.

Otra objeción que se plantea con frecuencia contra mi asimetría es ésta: ninguna falsación puede ser absolutamente segura, debido al hecho de que nunca podemos estar completamente seguros de que los enunciados básicos que aceptamos son verdaderos. Discutí plenamente esto en el capítulo de *L. I. C.* sobre la «base empírica» (capítulo V) y no creo que haya otra epistemología que lo haya tomado tanto en conside-

ración como la mía. (Los teóricos de la inducción, por ejemplo, nunca discuten de lleno este problema, a pesar del hecho de que sus teorías se derrumban si su base empírica resulta no ser firme.) Respecto a la afirmación de que este hecho refuta la asimetría entre la falsación y verificación, la situación es realmente muy simple. Tomemos un enunciado básico o un conjunto finito de enunciados básicos. Continúa siendo siempre una cuestión abierta si los enunciados son o no verdaderos: si los aceptamos como verdaderos, podemos haber cometido un error. Pero, *independientemente de si son verdaderos o falsos,* no puede derivarse de ellos una ley universal. Aunque estuviéramos seguros de que eran verdaderos, seguiría sin poder derivarse de ellos una ley universal.

Sin embargo, si suponemos que son verdaderos, una ley universal puede resultar falsada por ellos.

Por esta razón, la asimetría consiste en que un conjunto finito de enunciados básicos, *si son verdaderos,* puede falsar una ley universal; mientras que *bajo ninguna condición* podría verificar una ley universal: existe una condición bajo la cual el conjunto finito de enunciados básicos podría falsar una ley general, pero no existe ninguna condición bajo la cual pueda verificar una ley general.

Así, si aceptamos como verdadero el enunciado «este cisne de aquí es negro», entonces estamos obligados, por lógica, a admitir que hemos refutado la teoría universal «todos los cisnes son blancos»; y si aceptamos como verdadero en enunciado «este planeta está ahora más alejado del Sol de lo que estaba hace un mes», entonces estamos obligados, por lógica, a admitir que hemos refutado la teoría «todos los planetas se mueven describiendo circunferencias con el Sol como centro común». Ahora bien, es verdad —especialmente en el segundo caso— que podemos haber cometido un error al aceptar el enunciado singular en cuestión; y por esta razón, la falsación de la teoría *no es* «absolutamente segura». Pero *es* absolutamente seguro que *si* aceptamos cualquier enunciado singular («enunciado básico») que contradiga una teoría que habíamos aceptado, debemos haber cometido algún error, un error que debe ser corregido. Y es absolutamente seguro que si aceptamos un enunciado básico que contradiga la teoría que estamos contrastando, entonces estamos obligados a rechazar la teoría como falsada. Y es, por tanto, absolutamente seguro que (puesto que todo enunciado básico contradice algunas teorías) siempre que aceptamos *cualquier* enunciado básico, *algunas teorías* resultan implícitamente declaradas falsadas, *de modo que estamos comprometidos lógicamente a rechazarlas.* Pero *ninguna* teoría ha

sido verificada: no hay ninguna que estemos obligados a aceptar como verdadera. De ahí la asimetría.

De hecho, la asimetría es más fuerte incluso de lo que he indicado hasta ahora. Un principio tradicional del empirismo que yo acepto es que las teorías deben juzgarse a la luz de la evidencia observacional. Pero esto significa que al menos algunas veces tenemos que decidirnos a aceptar algún enunciado básico, aunque sólo sea provisionalmente, y después de muchas contrastaciones y deliberaciones. Y una vez que lo aceptamos, estamos lógicamente obligados, como hemos visto, a *rechazar* alguna teoría. No hay nada análogo a esto en lo que se refiere a la *aceptación de una teoría* o en lo que se refiere a su *verificación.*

Así, la relación lógica entre enunciados básicos y teorías y la incertidumbre de los enunciados básicos se refuerzan mutuamente en vez de suprimirse: *ambas operan en contra de la verificación; y ninguna opera unilateralmente contra la falsación.*

La cuestión de la incertidumbre de la base empírica está en un plano totalmente diferente de la cuestión de la relación lógica entre enunciados básicos y teorías. Tiene el mismo carácter del truismo según el cual siempre podemos errar (incluso en una demostración matemática); mientras que el carácter de la no-verificabilidad de las teorías es más parecido al de la advertencia, más interesante, de que nunca podemos tener suficientes premisas empíricas —suponiéndolas verdaderas— para establecer la verdad de una ley universal. Así pues, las dos cuestiones deben tratarse por separado. En resumen, es cierto —como resalté en *L. I. C.*— que ni siquiera las falsaciones son nunca absolutas, ni plenamente seguras; pero las razones para esta falta de seguridad son totalmente diferentes de las razones que hacen *imposible en principio,* y no simplemente algo dudosa, cualquier verificación de una teoría. El principio del propio empirismo entraña las dos: la asimetría y la posibilidad de falsar teorías.

Todo esto es tan trivial que no lo hubiera renunciado aquí si no fuera porque la existencia de una asimetría entre falsación y verificación ha sido negada constantemente por la mayoría de mis críticos positivistas.

Más grave es una objeción estrechamente relacionada con el problema del *contexto* y con el hecho de que mi criterio de demarcación es aplicable a *sistemas de teorías* más que a enunciados fuera de contexto. Esta objeción puede expresarse como sigue: Puede decirse que ninguna hipótesis singular es falsable, porque toda refutación de una conclusión puede tocar a una sola premisa del conjunto de todas las premisas usadas para derivar la conclusión refutada. La atribución de la falsedad a

una hipótesis determinada que pertenece a este conjunto de premisas es, por tanto, arriesgada, sobre todo si consideramos el gran número de suposiciones que entran en cada experimento.

Esta objeción pertenece a aquellos sobre quienes discutí en *L. I. C.* (secciones 16, último párrafo, sobre todo nota 2, y 18, 19, 20 y otros lugares). La respuesta es que, en efecto, sólo podemos falsar *sistemas de teorías* y que cualquier atribución de la falsedad a un enunciado determinado dentro de ese sistema es siempre sumamente inseguro.

Esto no afecta, desde luego, a la fundamental asimetría que he señalado. Pero me deja la tarea de explicar el hecho innegable de que algunas veces tenemos muchísimo éxito al atribuir a un hipótesis única la responsabilidad de la falsación de una teoría completa o de un sistema de teorías.

Muchos aspectos de nuestros procedimientos metodológicos reales puede entenderse como debidos a nuestros esfuerzos por lograr mayor éxito en esas atribuciones. Está, en primer lugar, la estructura en capas de nuestras teorías: las capas de profundidad, universalidad y precisión. Esta estructura nos permite distinguir entre las partes de nuestra teoría que corren más riesgos, que están más expuestas, y las otras que podemos, *relativamente hablando,* dar por sentadas cuando contrastamos una hipótesis expuesta. Esto explica el hecho de que contrastemos conscientemente, por regla general, cierta hipótesis escogida y tratemos al resto de las hipótesis implicadas en las contrastación, más o menos, como si no presentaran problemas, como una especie de «información básica»[35]. Esta información básica la *variamos* normalmente durante las contrastaciones, lo que tiende a neutralizar los errores que puedan estar implicados en ella.

Quizá el aspecto más importante de este procedimiento sea que siempre tratemos de descubrir cómo podríamos disponer *contrastaciones cruciales* entre la nueva hipótesis sometida a investigación —la que estamos tratando de contrastar— y algunas otras. Esto es consecuencia del hecho de que nuestras contrastaciones sean *intentos de refutación;* que están cocebidas *—concebidas a la luz de alguna hipótesis rival—* con el objeto refutar, si es posible, la teoría que queremos contrastar. Y siempre intentamos, en una contrastación crucial, hacer jugar el mismo papel a la información básica —en la medida de lo posible— con respecto a cada una de las hipótesis entre las que tratamos de forzar una decisión por medio de la contrastación crucial. (Duhem criticó los experimentos cruciales, mostrando

[35] [Para la idea de «información básica», véase *El desarrollo del conocimiento científico (C. & R.),* capítulo 10 y Addendum 3. *Ed.*]

que no podían establecer o demostrar una de las hipótesis en competición, como se suponía que hacían; pero, aunque trató de la refutación —señalando que su atribución a *una* hipótesis en vez de a otra era siempre arbitraria— nunca trató de la función que yo sostengo que es la de las contrastaciones cruciales: la de *refutar* una de las teorías en competición.)

Todo esto, claramente, no puede evitar en absoluto un error judicial: puede ocurrir que condenemos a una hipótesis inocente. Como he mostrado (especialmente en *L. I. C.*, secciones 19 y 20 y también 29), la aceptación de una refutación, o su atribución a una hipótesis en vez de a otra, siempre hay implicado un elemento de libre elección y de decisión.

Nuestros procedimientos científicos no están nunca enteramente basados en reglas; siempre hay suposiciones y corazonadas implicadas; no podemos eliminar de la ciencia el elemento de conjetura y riesgo.

¿Cómo podemos reducir este riesgo? Sólo tratando de calcular, lo mejor que podamos, las consecuencias de cada decisión —es decir, de cada ajuste de nuestra teoría— que parezca prometedora. (Tenemos que calcular, por así decirlo, todas las combinaciones prometedoras.) En este aspecto, la situación es similar a cualquier otro caso en el que tengamos que elaborar una nueva teoría: la decisión de atribuir la refutación de una teoría a una parte determinada de ella equivale a la adopción de una hipótesis; y el riesgo implicado es precisamente el mismo. Para hacerle frente necesitamos ingenio, osadía... y algo de suerte[36].

Por tanto, no hay procedimiento rutinario, no hay mecanismo automático, para resolver el problema de atribuir la falsación a ninguna parte determinada de un sistema de teorías, del mismo modo que no hay un procedimiento rutinario para idear nuevas teorías. El hecho de que no todo sea lógica en nuestra interminable búsqueda de la verdad no es, sin embargo, razón para que no usemos la lógica para iluminar lo mejor posible nuestra búsqueda, señalando dónde se desbaratan nuestros argumentos y hasta dónde alcanzan. La asimetría lógica fundamental que he descrito puede arrojar algo de luz, desde luego, sobre esta cuestión.

[*Todo esto trata de la *falsación empírica* y sus incertidum-

[36] *John W. N. Watkins ha señalado que la decisión sobre »dónde* estriba el error puede muy bien ser cuestión de una prolongada controversia crítica» y que eso puede ser lo mejor, puesto que es probable que esta controversia ayude a clarificar las consecuencias de los diversos cambios posibles en la teoría; véase su artículo «Epistemology and Politics», *Proc. Aristot. Society*, 1957, sobre todo el texto correspondiente a la última nota a pie de página (que se refiere a Duhem, *The Aim and Structure of Physical Theory*, capítulo VI).

bres. Debe distinguirse del criterio puramente *lógico* de *falsabilidad;* es decir, la *existencia* (*no* la verdad) de los falsadores potenciales de una teoría. No hay dificultades similares en relación con la *falsabilidad.* A la falsabilidad no le afectan los problemas que pueden afectar a las falsaciones empíricas.]

23. POR QUE INCLUSO LAS PSEUDO-CIENCIAS PUEDEN MUY BIEN TENER SENTIDO. PROGRAMAS METAFISICOS PARA LA CIENCIA

Mi criterio de demarcación —es decir, la contrastabilidad— le es necesario al científico del mismo modo que al filósofo en ciertas dificultades concretas. Separa las teorías que pueden discutirse seriamente en términos de experiencia. Advierte al científico de que hay otras teorías que no pueden discutirse de ese modo; y le llama la atención sobre el hecho de que esas otras teorías, al no ser contrastables, deben examinarse por otros métodos diferentes de contrastación. Si no encuentra otro medio de examinarlas críticamente, puede considerarse justificado para descartarlas. («La irrefutabilidad no es una virtud sino un vicio.» *Cf.* sección 7, *supra.*) Pero al hacerlo, siempre correrá un riesgo; porque a veces es posible aprender algo de verdadero interés incluso de una teoría pseudo-científica o de una metafísica.

Podemos considerar la astrología como ejemplo clásico de pseudo-ciencia. Su origen, junto con el de la astronomía, proviene de la creencia religiosa de que los planetas eran dioses (como lo afirmaba incluso el propio Platón). Esta creencia politeísta se abandonó en la astrología y en la astronomía, y ambas concordaron en que los planetas simplemente tenían los *nombres* de los dioses. Pero la astrología, aunque abandonó el politeísmo, no sólo continuó dando una significación mágica a los antiguos nombres divinos, sino que además siguió atribuyendo a los planetas poderes típicamente divinos que trató como «influencias» calculables. No es extraño que los aristotélicos y otros racionalistas rechazasen la astrología. Sin embargo, la rechazaron en parte por razones erróneas; y llevaron demasiado lejos su rechazo. La teoría lunar de las mareas, por ejemplo, fue, históricamente, hija del saber astrológico popular. Antes de su aceptación por Newton, la mayoría de los racionalistas la rechazaban como ejemplo de una superstición astrológica. Pero la teoría de Newton sobre la gravitación universal mostró no sólo que la luna podía influir en los «sucesos sublunares», sino que, además, algunos de los cuerpos celestres supra-lunares ejercían una influencia efectiva, un tirón gravitarorio, sobre la

tierra y, por tanto, sobre los sucesos sublunares, en contradicción con la teoría aristotélica. Así, Newton aceptó, con renuencia pero conscientemente, una doctrina que había sido rechazada por los mejores cerebros, incluido Galileo, como parte de una pseudo-ciencia desacreditada.

Esto muestra lo fácilmente que podemos perder una idea de la mayor importancia al rechazar por completo una teoría pseudo-científica.

Kepler, cuyas teorías eran una curiosa mezcla de la ciencia y astrología, proporciona un buen ejemplo de lo complicado que puede llegar a ser esto. A diferencia de Newton, que aceptó las ideas astrológicas sólo con renuencia, Kepler pertenecía a la tradición astrológica. Como Copérnico, Kepler pertenecia a la tradición platónico-pitagórica y creía en las «influencias» astrales, especialmente en las del Sol sobre los planetas. No obstante, era un astrónomo sumamente sutil y crítico de sí mismo: nunca se cansaba de someter sus hipótesis a contrastaciones ingeniosas y sumamente críticas, y analizaba las consecuencias a la luz de la mejor evidencia astronómica posible. Su actitud, de admirable autocrítica («Qué loco fui», escribió), le permitió hacer grandes contribuciones a la ciencia, a pesar del carácter fantástico de algunas de sus preciosas hipótesis[37].

Por cierto, los astrólogos siempre se han vanagloriado de que sus teorías se basan en un enorme número de verificaciones, en una evidencia inductiva abrumadora. Esta afirmación no ha sido nunca seriamente investigada ni desacreditada, y no veo por qué no podría ser verdad. Sin embargo, no tiene ningún interés ver con cuánta frecuencia ha resultado verificada la astrología; la cuestión es si se ha contrastado alguna vez con seriedad, con sincero deseo de falsarla.

Podemos dejar ahora la pseudo-ciencia de la astrología y ocuparnos de una teoría metafísica importantísima: el atomismo. El carácter metafísico de la «teoría corpuscular», en cualquier caso antes de Avogadro, es claro. No había posibilidad de refutarla. El hecho de que no se detectasen los corpúsculos, ni evidencia de ellos, podía explicarse siempre diciendo que eran demasiado pequeños para poder detectarlos. Sólo con una teoría que llevó a un cálculo del tamaño de las moléculas,

[37] Aquí vemos, como ocurre tan frecuentemente (por ejemplo, en el caso de Copérnico), que hipótesis importantes pueden tener su origen en ideas verdaderamente fantásticas: el origen no importa, siempre que la hipótesis sea contrastable. La frialdad de Galileo hacia Kepler y sus teorías es comprensible: Galileo pertenecía al campo racionalista (como la Iglesia Católica) y se oponía a la astrología. Esta es la razón, naturalmente, de que estuviese tan obsesionado con su teoría anti-astrológica, pero errónea, sobre las mareas y con su versión, ultrasimplificada, del copernicanismo.

quedó más o menos cerrada esta vía de escape, de modo que la refutación se hizo posible, en principio. (Las «verificaciones» fueron, en principio, posibles antes de esto: la invención del microscopio podría haber llevado al descubrimiento de moléculas visibles al microscopio.) Así, el atomismo se hizo contrastable tan pronto como se comprometió a un cálculo del tamaño de una molécula. Este ejemplo muestra que una teoría no contrastable —una teoría metafísica— puede ser desarrollada y fortalecida hasta que se convierte en contrastable. Pero si esto ocurre, parece que describirla como carente de sentido es un craso error; y es muy arriesgado rechazarla por completo, como hizo Mach.

Me he referido repetidamente al ejemplo del atomismo por ser enormemente característico e importante[38]. El atomismo antiguo era un sistema metafísico no sólo en el sentido de que no era contrastable, sino también en el sentido de que concebía el mundo en términos de una vasta generalización, de la mayor escala: «*No existe nada sino los átomos y el vacío*» (Leucipo, Demócrito)[39]. Sus dos conceptos fundamentales, átomos y vacío, eran inobservables y, por tanto, desconocidos, como señaló Demócrito, con lógica devastadora[40]. Así, el atomismo *explicaba lo conocido por lo desconocido.* Construía un mundo desconocido e invisible tras nuestro mundo conocido[41]. Y fue, precisamente por esa razón, persistentemente atacado por los positivistas (incluso después de 1905), por todos los inductivistas, desde Bacon hasta Mach y por todos los instrumentalistas, desde Berkeley hasta Duhem. Desde 1905 los positivistas han sido, es comprensible, más discretos en este punto. Sin embargo, nunca han explicado cómo puede ocurrir que un galimatías sin sentido se transustancie en algo con sentido. De hecho, el ejemplo del atomismo establece lo inadecuada que es la teoría de que la metafísica no es más que un galimatías sin sentido. Y establece lo inadecuada que es la política de hacer pequeños cambios subrepticios aquí y allá a la doctrina de la carencia de sentido, con la vana esperanza de salvarla[42].

Sea como fuere, el atomismo es un excelente ejemplo de una teoría metafísica incontrastable cuya influencia en la ciencia ha excedido la de muchas teorías científicas contrastables. Otra grandiosa teoría de este tipo fue la teoría de Descartes del

---

[38] *Cf.* secciones 4 y 85 de mi *L. I. C.*, y secciones 17 y 19, *supra.*

[39] Diels, *Vorsokratiker*, II, 6.ª ed., 1952, pp. 79 (15); 84 (10); 168 (6).

[40] *Cf.* Demócrito, Fragm. 117. (Diels, *op. cit.*, p. 168.)

[41] «Pero, en realidad, nada sabemos por haberlo visto; pues la verdad está oculta en lo profundo.» Demócrito, Fragm. 117. (Diels, *op. cit.*, p. 166; *cf.* el lema de este tomo.)

[42] *Cf.* sección 26 *infra.*

mundo como mecanismo de relojería (como me gusta llamarla a mí, porque se basa en la doctrina de que toda causación física era por empuje), o, como puede llamarse, el programa de Hobbes, Descartes y Boyle, de interpretar el mundo físico en términos de materia extensa en movimiento. Pero el último programa, y, hasta el momento, el más grande, fue el programa de Faraday, Maxwell, Einstein, de Broglie y Schrödinger de concebir el mundo —los átomos, así como el vacío— en términos de campos continuos[43]. Digo «fue» porque este maravilloso programa ha sido destruido por otros grandes físicos. (Véase el tomo III del *Post Scriptum*.)

Cada una de estas teorías metafísicas sirvió, antes de hacerse contrastable, como programa metafísico de investigación para la ciencia. Indicaba la dirección de nuestra investigación y el tipo de explicación que podía satisfacernos; y hacía posible algo parecido a una evaluación de la profundidad de una teoría. En biología, la teoría de la evolución, la teoría de la cédula y la teoría de la infección bacterial han tenido, todas ellas, papeles similares, al menos durante un tiempo. En psicología, el sensacionalismo, que puede tomar la forma de un tipo de atomismo psicológico (es decir, la teoría de que todas las experiencias están compuestas de elementos primarios no analizables, tales como, por ejemplo, los datos sensoriales), y el psicoanálisis deben mencionarse también como programas metafísicos de investigación.

Por importantes que hayan sido esos programas para la ciencia hay que distinguirlos de las teorías contrastables que el científico usa de modo diferente. De esos programas, él hace derivar su objetivo, lo que él consideraría una explicación satisfactoria, un descubrimiento real de lo que está «oculto en lo profundo». Aunque empíricamente irrefutables, esos programas metafísicos de investigación están abiertos a la discusión; pueden ser cambiados a la luz de las esperanzas que inspiran o de los desengaños de los que pueden considerarse responsables.

Naturalmente, hay también ideas metafísicas sin valor (o eso creo yo) para la ciencia. Algunos de los enunciados puramente existenciales discutidos en la sección 15 (véase también la 27) de *L. I. C.*, pertenecen a esta categoría. La aserción «existe una serpiente marina» no tiene especial interés para el científico mientras no se le añada alguna indicación que dé la clave de

[43] Pedí a J. Agassi que estudiase este tema y estoy en deuda con una obra suya que todavía no se ha publicado, por muchos detalles nuevos e interesantes sobre lo programas metafísicos de Bougle y de Faraday. [*Cf.* Joseph Agassi, «The Function of Interpretations in Physics», University of London Library, 1956; y el libro de Agassi, *Faraday as a Natural Philosopher*, 1971. *Ed.*]

posibles contrastaciones. No obstante, esas aserciones puramente existenciales han resultado a veces sugestivas, e incluso fuctíferas, en la historia de la ciencia, incluso aunque nunca llegaran a formar parte de ella. De hecho, pocas teorías metafísicas ejercieron mayor influencia en el desarrollo de la ciencia que la teoría puramente metafísica «existe una piedra filosofal (es decir, una sustancia que puede convertir los metales bajos en oro)», aunque no es falsable, nunca fue verificada y ahora ya no la cree nadie. El ejemplo de la piedra filosofal muestra, por cierto, que al menos en este caso, es el carácter existencial del enunciado y no la aparición de *términos* vagos o carentes de sentido lo que determina la incontrastabilidad: «metal bajo» y «oro» son suficientemente buenos para lo que suelen ser los *conceptos*.

CAPITULO III

METAFISICA: ¿CON SENTIDO O SIN SENTIDO?

> *El método correcto de la filosofía debería ser, en realidad, éste: No decir nada...: y luego, cuando alguien tratase de decir algo metafísico, mostrarle que no había dado significado a alguno de los signos de sus oraciones.*
>
> LUDWIG WITTGENSTEIN

Observación preliminar (1982). Esta cita, de uno de los últimos párrafos del *Tractatus* se convirtió en el programa de un movimiento mudial en filosofía. Su objetivo era mostrar su propia falta de sentido. ¡La metafísica es absurda! ¡La filosofía es un galimatías!

Ahora bien, a menudo me impaciento leyendo escritos filosóficos. Estoy bastante bien dispuesto a conceder que muchos de ellos son poco menos que un galiamatías: filosofía sin problemas auténticos. Así que comprendo la tendencia del *Tractatus* de Wittgenstein.

Pero, como veremos, las afirmaciones metafísicas pueden muy bien tener sentido y ser interesantes. Por «metafísica» entiendo aquí algo que es «incontrastable empíricamente». Wittgenstein quería decir «no plenamente verificable»[1]. Estas

[1] Según ciertas leyendas de moda, Wittgenstein nunca sostuvo el criterio de sentido según la verificabilidad: fue mal entendido y mal interpretado por Schlick y Waismann (que enunciaron el criterio de verificabilidad en *Erkenntnis* 1, 1930, pp. 228 ss.). Que la leyenda es falsa puede verse en el artículo de Schlick en *Die Naturwissenchaften* 19, 1931, pp. 145 ss.; especialmente p. 156, donde Schlick dice que una ley natural «puede ser modificada a la luz de una nueva experiencia» y, por tanto, no puede ser verificada nunca de forma concluyente o absoluta. En consecuencia, «no tiene el carácter lógico de un enunciado»: no es un enunciado sino (si se me permite usar la formulación de Ryle de una teoría de J. S. Mill) una licencia de inferencia. Schlick atribuye explícitamente esta teoría a una comunicación personal de Wittgenstein; y sin duda, obtuvo la aprobación de Wittgenstein antes de publicar los pasajes pertinentes de su artículo. Esto muestra que, en cualquier caso en 1931, Wittgenstein exigía de los «auténticos enunciados» («*echte Sätze*») que pudieran ser «concluyentemente» o «absolutamente» verificados. Todo esto puede ser comprobado mediante las publicaciones de Schlick que fueron escritas bajo la influencia

ideas y otras relacionadas con ellas se discutirán en este capítulo. Resulta que los dos programas que sugirió Wittgenstein son equivocados: 1) que las afirmaciones no verificables carecen de sentido porque violan las reglas gramaticales de la lengua: *Tractatus* 5.437 a 5.47321; 2) que carecen de sentido porque emplean palabras o expresiones aunque no se les haya dado un sentido: *Tractatus* 5.473 a 5.4733 y 6.53[2].

24. OBSERVACIONES LOGICAS SOBRE CONTRASTABILIDAD Y METAFISICA

En esta sección me propongo explicar, con ayuda de una serie de ejemplos, que los enunciados contrastables e incontrastables (o «metafísicos») pueden tener la misma forma lógica. Además, las mismas expresiones —palabras o símbolos— pueden aparecer en ellos (quizá en un orden ligeramente diferente) y que esas expresiones pueden tener precisamente el mismo significado. Así pues, ni la forma lógica de un enunciado, ni el tipo de expresión que aparece en él son suficientes para determinar si el enunciado es contrastable o incontrastable. Esta distinción, pues, no es entre enunciados bien formulados y pseudo-enunciados mal formulados (como suponía el programa positivista lógico). Sin embargo, no es una distinción que quede fuera del ámbito del análisis lógico.

Aparte de los enunciados francamente universales y los francamente existenciales, a menudo nos encontramos, tanto dentro como fuera de la ciencia, con enunciados con una estructura ligeramente más compleja. El enunciado que afirma la existencia de la piedra filosofal es un caso a propósito. Porque puede escribirse en la forma:

---

directa de sus discusiones con Wittgesntein y aprobadas por él. (Puedo añadir que esto me lo confirmó Schlick personalmente en una conversación en la que defendió apasionadamente sus teorías y las de Wittgenstein contra mis críticas. También me lo confirmó personalmente Waismann.)

[2] Los dos programas llegaron a ser, durante un tiempo considerable, la principal ocupación del Círculo de Viena y especialmente de su lógico más importante, Carnap. Véase su *Pseudo-problems in Philosophy (Schein-probleme in der Philosophie)*: «The Overthrow *(Überwindung)* of Metaphysics through the Logical Analysis of Language»; *The Logical Syntax of Language* (véase especialmente p. 278, donde se afirma que «el análisis lógico muestra» que los espurios enunciados de la metafísica son «pseudo-cláusulas»). El interés continúa estando en *Testability and Meaning* (donde mi criterio de demarcación según la contrastabilidad se malinterpreta como un criterio de significado). Por lo que yo sé, Carnap discutió algunos ejemplos de mala metafísica, pero nunca mostró «por medio del análisis lógico», como se proponía, que la no verificabilidad o no contrastabilidad sean siempre debidas a 1) o a 2). Como muestro en este capítulo, no lo son.

> «Existe un x tal que x es una sustancia y para todo y se cumple lo siguiente: si y es un metal bajo y se le añade o mezcla una pequeña cantidad de x a y, entonces y se convierte en oro.»

Podemos llamar a un enunciado de este tipo *enuncidado existencial-universal,* porque afirma la existencia de un x que se caracteriza por su relación con *todo* y de un cierto tipo.

Como tipo diferente de ejemplo, podemos considerar el igualmente incontrastable enunciado: «Todo suceso tiene una causa». Esto podría escribirse:

> «Para todo suceso x existe un y y un z tales que y sea una regularidad capaz de ser descrita por alguna ley universal (verdadera) u, y z sea un suceso (un conjunto de condiciones iniciales) que prececeda a x, y x sea predictible (deducible) de z en presencia de y (o de u)».

Podemos describir esto como un *enunciado universal-existencial.* No es contrastable porque no hay nada en él que nos sugiera cómo averiguar la existencia de la regularidad y (o de la ley universal u) y del suceso precedente z, y si no conseguimos encontrar un y y un z adecuados, no hay nada que nos indique si el fracaso se debe a nuestra incompetencia o al hecho de que y y z estaban bien escondidos o debido al hecho de que no existían tales y y z (es decir, que nuestro enunciado «todo suceso tiene una causa» era falso).

Llamé la atención sobre los enunciados de estas formas en la sección 66 de *L. I. C.* cuando traté del llamado «axioma de convergencia» o «axioma del límite» de von Mises[3].

Puesto que mi discusión dio lugar a malentendidos (*cf.* nota *2 a la sección 66 de *L. I. C.*) una discusión nueva y más exhaustiva parece apropiada.

Quiero llamar la atención tanto sobre el método de la discusión como sobre sus resultados. Ambos me parecen importantes y me parece que entrañan serias críticas a los métodos de moda: al método de construir lenguajes modelo por cuestiones de «precisión» y al método de análisis lingüístico o conceptual.

Mi aserción es que todas las expresiones o símbolos usados en dos enunciados, de los cuales uno es contrastable y el otro metafísico, pueden no sólo ser los mismos, sino que pueden

---

[3] El «axioma» de von Mises (que postula la existencia de un límite de la frecuencia relativa de la ocurrencia de una propiedad P en cualquier sucesión probabilística de sucesos o «colectivo») puede escribirse como un enunciado universal-existencial-universal-existencial-universal, de la forma siguiente: «*Para toda* sucesión probabilística, *existe* un número real x entre 0 y 1, llamado el límite de la frecuencia relativa, tal que *para toda* fracción dada y, por pequeña que sea, en la que $y > 0$, *existe* un número natural n_y tal que *para todo* número natural n (siempre que $n > n_y$), la frecuencia relativa, m/n, de m ocurrencias de la propiedad P hasta el enésimo suceso de la sucesión, no se desvía de x en más de y; es decir, $-y \leq x - (m/n) \leq +y$».

tener *precisamente el mismo significado.* Esto tiene una importancia considerable. Lo mismo ocurre con el hecho de que mi método haga posible afirmar que el significado es precisamente el mismo, aunque puede ocurrir que no sepamos lo que es «significado» y podemos ignorar el problema (si es que hay problema). El método que voy a adoptar es el siguiente: analizaré ciertas sencillas *conjeturas de matemática pura,* ciertos enunciados de teoría de los números. Como pertenecen a la teoría de los números, esas conjeturas *pueden* tener un carácter diferente de las conjeturas de las ciencias naturales, pero para nuestros propósitos, esas diferencias no importan. (Sigo creyendo que hay diferencias, aunque mi colega, el doctor I. Lakatos, ha puesto en duda esta creencia)[4].

Podemos considerar los números naturales 1, 2, 3, ... como constituyentes de todo nuestro infinito universo de discurso: son las cosas individuales en las que estamos interesados y sobre cuyas propiedades nos proponemos teorizar. Para los objetivos actuales, consideraremos que un enunciado es contrastable si, y sólo si, puede ser refutado (falsado) al examinar y determinar las propiedades de un número finito de esos individuos, e incontrastable (o metafísico) si no podemos refutarlo, por más números naturales que examinemos. (Se supone que no podemos examinar infinitamente muchos de ellos.)

Las únicas propiedades importantes de esos números que tomaré en consideración aquí son las propiedades «primo» y «compuesto»: un número natural se llama «primo» si, y sólo si, es mayor que 1 y no es divisible por ningún numero natural excepto por sí mismo y por 1; en otro caso, se llama «compuesto». (Así, el número 1 es el único número natural indivisible que es compuesto, y 2 es el único número natural par que es primo.)

Escojo para mi análisis las dos siguientes conjeturas muy simples y muy interesantes sobre números primos:

G) La conjetura de Goldbach.

H) La conjetura de los números primos gemelos.

Goldbach conjeturó que *todo* número par mayor que 2, es decir (*todo* número par que no sea primo), es la suma de dos primos. Esta conjetura puede expresarse también como sigue, puesto que todos los números pares mayores que 2 pueden escribirse en la forma $2 + 2x$.

> Para todo número natural x, se cumple los siguiente: $2 + 2x$ es la suma de dos números primos.
>
> Otra forma de expresar la misma conjetura es:

[4] [*Cf.* sus *Pruebas y refutaciones* (1976, èdición inglesa). *Ed.*]

G) Para todo número natural x existe al menos un número natural y tal que los dos números,

$$x + y \qquad \text{y} \qquad (2 + x) - y$$

sean, ambos, primos[5].

Es posible que la conjetura de Goldbach sea demostrada algún día. (Vinogradov ha demostrado un importante corolario.) Hoy no podemos probarla; pero hasta el momento no hemos encontrado un contra-ejemplo, un número natural x que no tenga la «propiedad de Goldbach», es decir, la propiedad que le atribuye G).

Vamos ahora a nuestro segundo ejemplo H), la llamada conjetura de los primos gemelos. Dos números primos que, como 3 y 5, ó 5 y 7, u 11 y 13, ó 17 y 19, ó 29 y 31, están separados sólo por *un* número natural (par), se llaman «primos gemelos». La conjetura de los primos gemelos enuncia que hay un número infinito de primos gemelos o, dicho de otro modo, que no existe un par máximo de primos gemelos. Puede expresarse de forma similar a la conjetura de Goldbach, como sigue:

H) Para todo número natural x existe al menos un número natural y tal que los dos números,

$$x + y \qquad \text{y} \qquad (2 + x) + y,$$

sean ambos primos[6].

---

[5] Notese que la suma de esos dos números es igual a $2 + 2x$. Escribiendo «*P (x)*» en lugar de «x es un primo» y usando el símbolo «*(x)*» con el significado de «para todo número natural x», «(E*y*)» con el significado de «existe un número natural y tal que» y «&» con el significado de la conjunción «y», podemos escribir la conjetura de Goldbach como sigue [nótese que la suma de $(x + y)$ y $(2 + x) - y$ es $2 + 2x$, es decir, un número par mayor que 2]:

G) $$(x)\,(\text{E}y)\,(P\,(x + y)\ \&\ P\,((2 + x) - y)).$$

Esta viene a ser la forma más sencilla de escribir la conjetura de Goldbach con este formalismo. También podemos definir (escribiendo «↔» en lugar de «si y sólo si») una propiedad *Gdb* (léase «goldbachiana») de números naturales tal que un número x tenga la propiedad *Gdb* si y sólo si, $2 + 2x$ es la suma de dos primos:

Def. 1. $$Gdb\,(x) \leftrightarrow (\text{E}y)\,(P\,(x + y)\ \&\ P\,((2 + x) - y)).$$

Con ayuda de este predicado «*Gdb*» podemos expresar ahora la conjetura de Goldbach por

G') $$(x)\ Gdb\ (x),$$

o, con palabras: «Todo número natural es goldbachiano».

[6] Usando los mismos símbolos que en la nota 5, la conjetura sobre los primos gemelos puede escribirse de una forma muy parecida a G):

H) $$(x)\,(\text{E}y)\,(P\,(x + y)\ \&\ P\,((2 + x) + y)).$$

Esta viene a ser, de nuevo, la forma más sencilla de escribir la conjetura sobre los números primos con este formalismo. También podemos definir una propiedad *Etp* (que debe leerse «excedido al menos por un par de primos

Nuestras dos conjeturas, G) y H), son en su expresión (aunque no en su contenido) sorprendentemente similares. La única diferencia es la aparición de un signo menos en G) en lugar en el que H) tiene su signo más. Aparte de eso, son exactamente iguales. Además, H) usa *sólo* expresiones que también aparecen en G); incluso el signo más aparece en G), aunque sólo una vez. Por tanto, está claro que si G) «tiene sentido» en cualquier sentido más o menos aceptable, H) debe «tener sentido» también, precisamente en el mismo sentido.

Sin embargo, si nos interesa la *contrastabilidad*, esta pequeña diferencia es *toda* la diferencia: porque mientras que G) es contrastable (falsable), H) no lo es. Podemos decidir, de cualquier número x dado si x tiene la propiedad de Goldbach; pero no si a x le exceden unos números primos gemelos.

Para contrastar G) para cualquier número x dado, es suficiente considerar aquellos números y que sean menores que x. (Sólo en el caso de que $x = 1$, consideremos que $x = y$; en todos los demás casos comprobados y los y entre 1 y $x - 1$, ambos inclusive.) Cada uno de los números para los que $x + y$ es primo, se comprueba entonces para ver si $2 + x - y$ es o no primo. Si uno de los números y hace tanto a $x + y$ como a $2 + x - y$ primos, entonces x tiene la propiedad de Goldbach; si no es así, no la tiene. De este modo, podemos decidir en la mayoría de los pasos $x - 1$ de este proceso de comprobación si algún x dado es o no goldbachiano.

Por el contrario, contrastar H) para algún número x dado implica comprobar todos los números que exceden a x en y, para $y = 1$, $y = 2$,... para ver si encontramos un y tal que $x + y$ y $2 + x + y$ sean ambos primos. Naturalmente, si encontramos, para algún x dado un y de este tipo, entonces habremos corroborado H) para ese determinado número x dado. Pero si no

---

gemelos») de los números naturales, tal que el número x tenga la propiedad *Etp* si, y sólo si, es menor que algún par de primos gemelos:

Def. 2. $Etp\ (x) \leftrightarrow (Ey)\ (P\ (x + y)\ \&\ P\ ((2 + x) + y))$.

Con ayuda del nuevo predicado «*Etp*» podemos expresar ahora la conjetura de los primos gemelos por

H') $(x)\ Etp\ (x)$.

o, con palabras: «Todo número natural es excedido por algún par de primos gemelos».

Por cierto, podemos atribuir la propiedad de «ser primos gemelos» o «*Twp*» a uno —digamos que al menor— de un par de primos gemelos; es decir, podemos definir

Def. 3. $Twp\ (x) \leftrightarrow P\ (x)\ \&\ P\ (2 + x)$

que nos permitiría sustituir Def. 2 por

Def. 2' $Etp\ (x) \leftrightarrow (Ey)\ Twp\ (x + y)$

encontráramos un y satisfactorio hasta, digamos, $x + y = 2x$, ó $x + y = x^2$, entonces tendríamos que seguir comprobando con números y cada vez mayores. E incluso aunque nunca encontremos un y satisfactorio, eso no significa que hayamos falsado H) para el número x dado; porque siempre es posible que un número y mayor pueda surtir efecto y satisfacer H).

Así, G) puede ser refutada por medio de un examen, es decir, al encontrar un ejemplo que la refute, un x que no la satisfaga. Pero H) no puede ser refutada de la misma forma, puesto que nunca podemos estar seguros, por medio de un examen, de que un x dado no satisface H). Esto significa que G) es falsable y que H) es infalsable. Naturalmente, ni G) ni H) son verificables.

Es interesante advertir, sin embargo, que el carácter de H) cambiaría totalmente si un día lográsemos demostrar un teorema que afirmase, por ejemplo, que si a x le excede al menos un par de números primos gemelos, entonces el menor de esos pares debe situarse entre x y un número calculable mayor que x. Por ejemplo algún día podemos lograr demostrar la siguiente conjetura falsable:

> H′) Para todo número natural x se cumple lo siguiente: siempre que $x > 1$, y siempre que á x le exceda al menos un par de primos gemelos, debe haber al menos un par de primos gemelos entre x y $2 + 2x$; o en otras palabras, debe haber al menos un y que sea el menor de un par de primos gemelos de tal forma que
>
> $$x < y < 2x$$

La conjetura H′) supone, entre otros enunciados, el siguiente: para todo par de primos gemelos, x y $x + 2$ —siempre que sean excedidos por algún par de primos gemelos y siempre que $x \geqslant 5$— existe un par de números primos gemelos mayores tales que el gemelo menor está entre $x + 6$ y $2x - 1$, ambos límites inclusive.

Si logramos demostrar la cojetura H′) (o quizá una más debil, por ejemplo, con $x < y < x^2$, o con $x < y < 2^x$, en lugar de $x < y < 2x$), entonces la conjetura de los primos gemelos H) se convertiría en contrastable (en el sentido cuasi-empírico discutido aquí). Porque una prueba de H′) haría posible en principio establecer cualquier número dado x, mediante el examen de un número finito de números diferentes, que no hay números primos gemelos mayores que x; o dicho de otro modo, que x falsa la conjetura de los primos gemelos.

Es interesante advertir que H′) es, a su vez, una conjetura falsable: podría ser falsada, por ejemplo, si encontrásemos dos pares de primos gemelos *consecutivos*, a, $a + 2$ y b, $b + 2$, tales que $b > 2a$.

No obstante, H^{f}), aunque falsable, no es sistemáticamente contrastable de la misma manera que, por ejemplo, G). Porque si, después de cierto número *a*, no encontramos otro primo gemelo, aunque hayamos comprobado todos los números hasta, digamos, 3*a* o 4*a*, entonces no sabemos si *a* es un ejemplo refutador de H^{f}) o no: 4*a* + 1 puede ser el primero de un par de primos gemelos [en cuyo caso, H^{f}) es falsa], o puede no haber ya ningún par más de primos gemelos [en cuyo caso, H^{f}) puede ser verdadera].

Hemos visto que H^{f}) es falsable; y en presencia de H^{f}), H) se hace también falsable. Por tanto, podemos decir que el carácter metafísico de H) es *relativo*, relativo con respecto al estado actual de nuestro conocimiento (*no* con respecto al «sistema de lenguaje» utilizado). Pues, aunque H) no es contrastable actualmente y es metafísica [el descubrimiento de una prueba de H^{f})] la haría contrastable. Además, un sistema de conjeturas que esté compuesto de H^{f}) y de H) constituye un sistema teórico contrastable.

Se nos plantea la cuestión de si podríamos sustituir, en este sistema teórico, H^{f}) por una conjetura incontrastable o metafísica [a la que podemos llamar H^{m})], y retener la contrastabilidad de la *conjunción* de ambas, es decir, de H) y H^{m}). La respuesta es: sí, y es muy fácil. H^{m}) puede ser, por ejemplo, la conjetura siguiente:

> H^{m}) Si a todo número le excede algún par de primos gemelos, entonces, para todo número *x*, si *x* y *x* + 2 son números primos gemelos, existe al menos un par (mayor) de primos gemelos, cuyo gemelo menor se halla entre *x* y 2*x*.

Aunque H^{f}) y H^{m}) difieren en su forma lógica, la diferencia en la fraseología de ambas es mínima y ciertamente las dos emplean las mismas expresiones o «conceptos». Esto puede verse claramente si escribimos H^{f}) como sigue.

> H^{f}) Para todo número *x* se cumple lo siguiente: si a *x* le excede algún par de primos gemelos, entonces, si *x* y *x* + 2 son primos gemelos, existe como nínimo un par (mayor) de primos gemelos cuyo gemelo menor se halla entre *x* y 2*x*.

H^{m}) se obtiene, de modo muy sencillo, debilitando a H^{f}) lo suficiente como para hacerla incontrastable. Sin embargo, aunque H^{f}) es contrastable y H^{m}) es metafísica, la conjunción de H) y H^{f}) tiene exactamente la misma fuerza lógica o el mismo contenido que la conjunción de H) y H^{m}), como puede verse fácilmente[7].

---

[7] Para que esto quede claro escribimos

H^{f}) $(x)\,(Etp\,(x) \rightarrow (Twp\,(x) \rightarrow (\mathrm{E}y)\,(Twp\,(y)\ \&\ x < y < 2x)))$.

Nuestros ejemplos G), H), H^f) y H^m) pueden utilizarse para mostrar una serie de puntos importantes.

1) La diferencia entre enunciados contrastables y enunciados metafísicos no depende necesariamente de su forma lógica[8], sino más bien de la información que transmiten, algo que puede expresarse de las formas más distintas y que no depende (si es expresable adecuadamente en algún lenguaje) del lenguaje escogido: si no es contrastable en el lenguaje L_1, puede serlo en el lenguaje L_2.

2) La diferencia no depende, ciertamente, de la aparición, en el enunciado metafísico de símbolos o expresiones carentes de sentido o mal definidos. Sólo aparecen en H) o H^m) los símbolos que aparecen en G) o H^f) respectivamente. En G) aparece un símbolo «—» que *no* aparece en H); pero es obvio que este hecho —que podría ser concebible que convirtiera a G) en carente de sentido— no puede hacer carente de sentido a H) si G) tiene sentido.

3) El carácter metafísico de H) y H^m) es relativo con respecto a nuestras otras conjeturas: en presencia de H^f), H) se hace contrastable. También, en presencia de H), H^m) se hace contrastable (y viceversa).

4) El carácter metafísico de un enunciado puede ser consecuencia de su debilidad lógica, de su falta de contenido: H^m) se obtiene a partir de H^f), al debilitarla (al disminuir su contenido).

Del mismo modo, podemos obtener un enunciado metafí-

H^m) $(x)\,(E!p\,(x)) \rightarrow (x)\,(Twp\,(x) \rightarrow (Ey)\,(Twp\,(y)\ \&\ x < y < 2x))$.

Quedará claro que H^m) se sigue de H^f) de acuerdo con el principio lógico, bien conocido, $((x)\,(A\,(x) \rightarrow B\,(x))) \rightarrow ((x)\,A\,(x) \rightarrow (x)\,B\,(x))$. También está claro que H), es decir, *(x) E!p (x), junto con* H^m), obtenemos H^f); así, la conjunción de H) y H^f) tiene precisamente la misma fuerza que la conjunción de H) y H^m).

[8] Podemos, desde luego, construir formalismos a los que puedan ser traducidos G) y H) y con respecto a los cuales pueda decirse que su diferencia es una diferencia de forma lógica. Por ejemplo, podemos estipular que el símbolo «(E*x*)» puede ir seguido de un subíndice que enuncie un límite superior para la variable *x*. En este caso, G) y H) puede convertirse, por ejemplo, en

G') $(x)\,(Ey)_{x-1}\,(x > 1 \rightarrow P\,(x + y)\ \&\ P\,((2 + x) — y))$.

H') $(x)\,(Ey)\,(x > 1 \rightarrow P\,(x + y)\ \&\ P\,((2 + x) + y))$.

La diferencia entre ellos quedaría expuesta por el hecho de que el operador «(E*x*)» está limitado en G) pero no en H), lo que, podría decirse, supone una diferencia en sus respectivas formas lógicas; y podemos, naturalmente, excluir a H) como carente de sentido al estipular que todos los operadores existenciales tengan que estar limitados (por alguna constante finita o por algún término diferente del que aparece en el propio operador). Pero lo cierto sigue siendo que podemos expresar ambas conjeturas de forma perfectamente adecuada en lenguajes en los que sus respectivas formas lógicas sean idénticas.

sico debilitando á G). Por ejemplo, podemos lograr, debilitando a G), la conjetura infalsable de que *casi todos* los números naturales (digamos que todos, excepto un conjunto finito) son goldbachianos (una conjetura estrechamente relacionada con la que Vinogradov ha demostrado). O podemos obtener la conjetura, aún más débil, de que para todo número x, existe un bloque de al menos x números goldbachianos consecutivos. Esas dos conjeturas son, obviamente, consecuencias incontrastables de G). (Una consecuencia aún más débil —la que existen infinitos números goldbachianos— es demostrable, puesto que es una consecuencia inmediata del teorema, demostrado por Euclides, de que existen infinitos números primos.)

5) El carácter metafísico de un enunciado no es una *mera* cuestión de debilidad lógica o de falta de contenido lógico. Aunque un enunciado (no tautológico) cuyo contenido *lógico* sea demasiado reducido no puede tener ningún contenido empírico y, por tanto, tiene que se metafísico, lo contrario no es cierto: un enunciado puede ser metafísico (lo·que significa que su *contenido contrastable* es cero) aunque puede tener un gran *contenido lógico*. Este hecho se establecerá en el próximo punto, 6)[8a].

6) Podemos ilustrar todo esto con ayuda de enunciados que empleen no sólo términos puramente matemáticos, sino también términos empíricos. El ejemplo más sencillo, quizá, se obtiene mediante la sustitución de algunos de nuestros términos matemáticos por términos empíricos.

Así, tomando una sucesión de tiradas con un penique como nuestro universo de discurso, podemos sustituir «x es primo» por «en la tirada número x sale cara». Obtenemos así, de nuestras dos conjeturas, G) y H), las siguientes *conjeturas sobre nuestra sucesión de tiradas del penique:*

[8a] [*(Añadido en 1982.) Es posible hacer aún más parecidos a G) y H). Podemos utilizar el símbolo $/x/$ para el «valor absoluto» de x; es decir $/x/ = x$ si, y sólo si, $x \geq 0$; en otro caso, $/x/ = -x$ (y, por tanto, positivo). De modo que podemos escribir:

F) $(x)\,(\mathrm{E}y)\,(P\,(/x/ + /y/)\ \&\ P\,(2 + /x/ + y))$.

Esta fórmula enuncia G) cuando permitimos a las variables x e y extenderse a los números enteros negativos $x, y < 0$; y la misma fórmula F) enuncia H) cuando dejamos que las variables x e y se extiendan a los enteros positivos, $x, y > 0$; y enuncia

I) $(x)\,(Gdb\,(x)$ v $Etp\,(x))$

cuando permitimos que x e y se extiendan a los enteros positivos y negativos excepto 0: $x, y \neq 0$.

Es interesante que aunque esta extensión incluye a las otras, I) es, con mucho, la más débil de las tres conjeturas G), H) e I). Es mucho más débil incluso que la disyunción de G) y de H).

GP) Para todo número natural x, existe al menos un número natural y tal que los dos números

$$x + y \qquad \text{y} \qquad (2 + x) - y$$

sean ambos números de tiradas en las que salgan caras.

HP) Para todo número natural x existe al menos un número natural y tal que ambos números

$$x + y \qquad \text{y} \qquad (2 + x) + y$$

sean ambos números de tiradas en las que salgan caras.

Ahora bien, G) es una conjetura muy dudosa sobre todos los números x, puesto que el hecho de que se cumpla para un x pequeño es cuestión de pura suerte; aunque al aumentar x se hace cada vez más probable que cualquier ejemplo x se conformará a GP).

Pero HP) es «casi seguramente verdadera» para todo x y para toda serie de tiradas del penique, porque solamente afirma que de vez en cuando encontraremos, si continuamos tirando, dos tiradas en las que salgan caras separadas solamente por *una* tirada (en la que no importa si salen caras o cruces). Esta es una conjetura muy débil: dada la información de que la probabilidad de que salgan caras es de 1/2, y de que las tiradas son independientes, el contenido lógico de HP) resulta ser cero[9].

De ahí se sigue que el contenido lógico de la negación de HP), dada la misma información, es 1. Esto ocurre porque la negación de HP) es «casi seguramente falsa», con la información dada: la negación de HP) «casi contradice» esta información[10]. No obstante, la negación de HP) no tiene contenido *empírico:* no es contrastable, precisamente porque afirma *la mera existencia* de un número de tiradas después del cual no habrá más tiradas gemelas de caras (separadas precisamente por una tirada). Por tanto, no será refutada, incluso aunque encontremos tales tiradas gemelas una y otra vez, después de cada número que examinemos.

El ejemplo es importante. Muestra la existencia de enunciados «puramente metafísicos» (es decir, de enunciados incontras-

---

[9] Supongo aquí que el contenido lógico de a dado b es complementario a la probabilidad de a dado b; es decir,

$$Ct\,(a,b) = 1 - p\,(a,b)$$

[10] En la teoría de von Mises, la negación HP) contradice de hecho la información; porque en esta teoría, el hecho de que una ocurrencia de tiradas gemelas de caras separadas exactamente por *una* tirada, tenga la probabilidad de 1/4, entraña que las ocurrencias de este tipo deben suceder una y otra vez. Así, en esta teoría, HP) se sigue de hecho de la información «dada» mientras que en la teoría clásica de la probabilidad y en las teorías que tratan de obtener medidas teóricas de la probabilidad, no se sigue (aunque «casi se sigue»). (Véase *infra*, secciones 22 y ss.)

tables cuyas negaciones tampoco son contrastables o, dicho de otro modo, enunciados que no son ni falsables ni verificables). Además muestra que podemos tener razones —quizá incluso razones empíricas— para aceptar un enunciado, aun «completamente metafísico» y rechazar así, en principio, su negación. (Esto, desde luego, es una cosa que ocurre normalmente con los enunciados metafísicos unilateralmente verificables, tales como «existe al menos un cisne que no es blanco»; aunque es cuestión de gustos el llamar a este enunciado metafísico o empírico, en la medida en que puede considerarse no como «aislado», sino como parte de un sistema empírico. Naturalmente, si un enunciado incontrastable *añade* al contenido lógico de un sistema contrastable —como en el caso de la conjunción de H) y H^m), o de sus contrapartidas empíricas— ya no lo llamamos metafísico cuando lo consideramos parte de dicho sistema.)

No hay necesidad de formular explícitamente las conjeturas sobre tiradas de monedas que son contrapartida de H^f) y H^m). En cambio, puede merecer la pena mencionar que podemos construir otras contrapartidas muy parecidas (aunque no muy interesantes) de nuestras conjeturas matemáticas en términos empíricos. Podemos escoger, por ejemplo, conjeturas sobre la actividad de las manchas solares e intepretar nuestros números naturales *x* como números de años (contando a partir de un momento cero) y sustituir nuestro predicado «primo» por «un año de poca actividad en las manchas solares».

7) Tanto el enunciado contrastable G) como el enunciado metafísico H) dan lugar, como hemos visto, a la definición de un predicado característico —los predicados «goldbachiano» y «excedido por algún par de primos gemelos»—. El primero de esos predicados puede llamarse «completamente decidible» o «normal» u «ordinario» o «recurrente», porque podemos decidir, inspeccionando un número finito de casos si un número dado *n* es o no goldbachiano. Es, en este aspecto, muy similar a muchos predicados empíricos totalmente ordinarios [véase 8) *infra*]. El segundo predicado, «excedido por algún par de primos gemelos» puede ser «decidible unilateralmente» o «seminormal» o «semi-ordinario» o «recurrentemente enumerable», porque es verificable pero no falsable: dado un número *n*, *podemos* hallar un par de números primos gemelos mayores que *n* y verificar así el enunciado «*n* es excedido por algunos primos gemelos»; pero aunque no encontremos un par de primos gemelos que lo excedan, el enunciado no resulta falsado.

Vemos aquí el esqueleto de una clasificación en categorías (una clasificación que podría refinarse mucho si así lo deseáramos):

a) Predicados «ordinarios» que, si se aplican a individuos singulares (por ejemplo a números individuales), dan enunciados singulares *completamente decidibles.*

b) Predicados «semi-ordinarios» que, si se aplican del mismo modo, dan enunciados singulares *unilateralmente verificables* (es decir, infalsables). (Estos predicados pueden llamarse «predicados metafísicos».)

c) Predicados «semi-ordinarios» que dan enunciados singulares *unilateralmente falsables.*

d) Predicados «extraordinarios» que dan enunciados singulares que no son *ni verificables ni falsables.* (Estos predicados pueden llamarse también «predicados completamente metafísicos».)

Veremos que el complemento o negación de un predicado de la categoría *a)* pertenece también a la categoría *a)*, y que la negación de un predicado de la categoría *d)* pertenece también a *d)*; pero la negación de un predicado de la categoría *b)* pertenece a *c)* y viceversa.

8) Algunos ejemplos y métodos de construcción de predicados (tanto empíricos como matemáticos) de las cuatro categorías están implícitos en nuestra discusión.

a) Como ejemplos de predicados «ordinarios» o «completamente decidibles» podemos tomar los predicados «primos» o «goldbachiano» o «rojo» o «tirada en la que sale cara» y, naturalmente, los complementos o negaciones de esos predicados.

Mientras que los *enunciados singulares* que se obtienen de esos predicados son completamente decidibles, los *enunciados universales* son, en un universo infinito, unilateralmente falsables y los *enunciados existenciales,* unilateralmente verificables.

b) Podemos construir ahora algunos predicados unilateralmente verificables a partir de los de categoría *a)*; por ejemplo, el predicado «excedido por algún número goldbachiano» y el predicado «excedido por algún número no goldbachiano» son ambos unilateralmente verificables (y también lo es el predicado «excedido por algún par de primos gemelos»).

Respecto a «rojo», podemos suponer un universo de coches y podemos llamar a un coche *c* «seguido de otro coche» si *c* es de una fecha anterior o tiene un número de chasis más bajo. El predicado «seguido por un coche pintado de rojo» será entonces un predicado verificable e infalsable mientras continúe la producción de coches. Observaciones análogas son válidas para «seguido por un año de poca actividad en las manchas solares» o «sucedido por cinco tiradas consecutivas de caras», etc.

Estos predicados dan lugar a enunciados singulares unilateralmente verificables y a enunciados existenciales unilateralmente verificables; pero los enunciados universales no son

ni verificables ni falsables: son completamente metafísicos, como H).

c) Los predicados unilateralmente falsables pueden construirse de la misma forma. Por ejemplo, «seguido sólo por números goldbachianos» o «seguido sólo por coches pintados de rojo» da lugar, claramente, a enunciados singulares que son unilateralmente falsables. [También puede hacerse referencia aquí a los complementos de predicados de la categoría *b)* y viceversa.] Los enunciados universales son también unilateralmente falsables, pero los enunciados existenciales son completamente indecidibles.

d) Finalmente, podemos construir predicados completamente indecidibles a partir de las categorías *b)* o *c)* mediante una especie de universalización de los de *b)* y mediante una especie de particularización (existencialización) de los de *c)*. Obtenemos así dos subcategorías, d_b) y d_c), pero hay otras muchas pertenecientes a *d)*. Por ejemplo, el predicado (demostrablemente universal) «excedido sólo por números que a su vez son excedidos, cada uno de ellos, por algún número goldbachiano» es un predicado completamente metafísico de la subcategoría d_b), mientras que «excedido por algún número que a su vez sea excedido sólo por números goldbachianos» pertenece a d_c). Los otros predicados en *b)* y en *c)* dan predicados análogos tales como «seguido sólo por coches, cada uno de los cuales es seguido, a su vez, por un coche pintado de rojo» d_b), o «seguido por algún coche que, a su vez, es seguido sólo por coches pintados de rojo» d_c).

Los predicados de d_b) y d_c), construidos según este esquema, dan lugar a enunciados singulares, universales y existenciales que son todos completamente incontrastables o «completamente metafísicos».

Pero ninguno de esos predicados y ninguno de esos enunciados debe describirse como carente de significado en ningún sentido. Por el contrario, los predicados están bien definidos en términos de conceptos obviamente significativos; y los enunciados están gramaticalmente bien construidos[11]. Además, puede

[11] Para mostrar esto en un lenguaje simbólico típico, podemos empezar con el predicado, completamente decidible, «*A*» («rojo» o «goldbachiano» o «no-goldbachiano», etc.) y la relación «<» («es excedido por»). Definimos

$$B\,(x) \leftrightarrow (\mathrm{E}y)\,(x < y\ \&\ A\,(y))$$

$$C\,(x) \leftrightarrow (y)\,(x < y \rightarrow A\,(y))$$

$$D_B\,(x) \leftrightarrow (y)\,(x < y \rightarrow B\,(y))$$

$$D_C\,(x) \leftrightarrow (\mathrm{E}y)\,(x < y\ \&\ C\,(y))$$

demostrarse que algunos de los enunciados son verdaderos y puede demostrarse que otros son falsos.

Por ejemplo, «excedido por algún número goldbachiano», que pertenece a *b)*, es un predicado demostrablemente universal. Es decir, puede demostrarse que es válido para todo número *n* (simplemente porque no hay número primo máximo y, por tanto, no hay suma máxima de dos números primos). Lo mismo ocurre con «excedido sólo por números que son excedidos por algunos números goldbachianos», que pertenece a d_b) y es, por consiguiente, «completamente metafísico». De hecho, todos los enunciados singulares, universales y existenciales que se obtienen de los predicados de la categoría *d)* formados de la manera que se ha descrito aquí, son consecuencias lógicamente débiles de enunciados de la categoría *a)*. Por ejemplo, «excedido por algún número que sea excedido sólo por números goldbachianos» d_c), es una forma ligeramente más débil, desde el punto de vista lógico, de «excedido sólo por números goldbachianos» *c)*[12].

Sería difícil para cualquiera describir estos ejemplos como carentes de sentido (o «cognoscitivamente carentes de sentido»); sobre todo para aquellos filósofos que arguyen, correctamente, en favor de la significatividad de los enunciados puramente existenciales y que creen, equivocadamente, que con ello pueden mostrar lo inadecuado de mi criterio de demarcación (que de hecho, caracteriza a los enunciados existenciales aislados e incontrastables y «metafísicos», aunque no de carentes de sentido, naturalmente).

El predicado «excedido por algún par de números primos gemelos» que se basa en H) es un predicado incontrastable; sin embargo, en presencia de H′) se hace contrastable.

Así, *la posición de un predicado* con respecto a la contrastabilidad puede cambiar con el avance de nuestro conocimiento; precisamente del mismo modo que la posición de un enunciado como H) con respecto a la contrastabilidad puede cambiar con el avance de nuestro conocimiento. Pero el *sentido* (lo que quiera que esto sea)[13] del predicado no necesita cambiar al mismo tiempo. En nuestro ejemplo, no cambia, ciertamente, puesto que su sentido está plenamente definido en todos los

---

Estas parecen ser las definiciones más sencillas que producen los resultados deseados; pero hay, naturalmente, muchas otras formas de lograr resultados similares.

[12] [El texto es defectuoso en este lugar. Faltan varias líneas con otros ejemplos y no pudieron ser reconstruidas. *Ed.*]

[13] No sé *qué es* el «significado» de una expresión (y no soy partidario de las preguntas del tipo «¿qué es?») pero supongo que una definición (matemática) puede al menos ayudar a determinarlo.

enunciados, en términos de los predicados «número natural», «primo», «mayor» y de algunos operadores lógicos, etc. Observaciones análogas son válidas para el predicado «excedido por tiradas de caras separadas exactamente por una tirada».

Este resultado tiene un interés considerable; al menos para mí, resuelve un problema que me ha inquietado durante muchos años. El problema es si el significado de nuestros términos teóricos cambia *siempre* o no, cuando debido al progreso de la ciencia, cambie su posición con respecto a la contrastabilidad. Que el significado de nuestros términos teóricos cambia con mucha frecuencia en el curso del progreso científico está claro: no tenemos más que pensar en términos como «átomo», «planeta», «luz», «movimiento», «espacio» o «fuerza». Influido por estos ejemplos, llegué a dudar de que existiera algo como la «invariación de significados» (el término es de Paul Feyerabend)[14], en una ciencia en crecimiento; y más en especial dudaba si el significado de un término teórico no estaría abocado a cambiar siempre que el avance de la ciencia afectase a su posición con respecto a la contrastabilidad.

Sin embargo, esas dudas son infundadas, como lo muestra nuestra discusión. Aunque el significado de nuestros términos cambie al cambiar nuestras teorías, tal cambio no tiene que seguir necesariamente también a un cambio de posición con respecto a la contrastabilidad. La contrastabilidad y el significado están, naturalmente, relacionados: si una teoría es contrastable o no dependerá en parte, del significado de los términos utilizados (y, en parte, del estado de nuestro conocimiento). Pero no hay dependencia en la otra dirección: ni siquiera cambios espectaculares en la contrastabilidad necesitan afectar al significado en cuanto tal.

Estos resultados ofrecen una solución de lo que fue para mí un problema abierto durante mucho tiempo. Cuando comprendí la solución, pensé también que quizá mis dudas fuesen un síntoma de que no estaba totalmente libre de la influencia del pensamiento positivista, aunque había argüido hacía mucho tiempo contra la teoría de que el significado de un término estaba determinado por el método de verificar (o incluso de contrastar) el enunciado en el que aparecía el término en cuestión. Nunca había estado realmente interesado en el problema del significado de los términos, y había aceptado acríticamente y desde luego, sin pensarlo, la concepción algo más amplia de que el significado de un término estaba determinado por su

[14] [*Cf.* sus «Problems of Empiricism» en Rober G. Colodny, ed., *Beyond the Edge of Certainty*, 1965, p. 164. Véase Popper, *L. I. C.*, sección 38, nota *3. *Ed.*]

uso. Puesto que el uso científico de un término resultaba influenciado con frecuencia por su posición respecto a la contrastabilidad, parecía que los cambios en esta posición podrían influir en el significado del término en cuestión.

Sólo como resultado del análisis del que trato en esta sección, advertí que todo esto era erróneo; que no es el «uso» de un término lo que determina su significado, sino, más estrictamente, *las relaciones que suponemos* válidas *entre las entidades* denotada por el término y otras entidades. Aquí, como siempre, el análisis lingüístico es equivocado. Nuestro uso (o sus reglas) pueden ser síntoma de nuestros supuestos teóricos —puede revelar esos supuestos— pero no es idéntico a ellos. Y aunque nuestros supuestos son, a menudo, implícitos y no son fáciles de formular explícitamente, pueden ser formulados explícitamente como *conjeturas sobre cosas*. El uso, por otro lado, o las reglas de uso, pueden formularse, en el mejor de los casos, como fórmulas que describen nuestro propio comportamiento.

(Por ejemplo, puede decirse que los axiomas de Hilbert dan el significado de la palabra «entre». Pero esos axiomas son suposiciones sobre cosas, y sobre relaciones especiales entre cosas. No son reglas que establezcan hábitos del habla.)

El hecho de que el significado de nuestros términos no dependa, en general, de su posición respecto a la contrastabilidad destruye, creo, otra teoría muy popular sobre el modo en que los términos teóricos adquieren significado empírico. Estoy refiriéndome a la teoría inductivista según la cual es a través de su contacto (indirecto) con la observación y el expremiento como se confiere significado empírico a los términos teóricos. (La teoría operacionalista del significado es una variante de este enfoque inductivista.) La idea intuitiva que subyace bajo esta concepción puede expresarse como sigue: los términos no teóricos o términos empíricos de carácter básico denotan «observables» y, por tanto, «tienen significado»; los términos teóricos deben, si es posible, definirse por medio de (o ser «constituidos» por o «reducidos» a) esos términos empíricos básicos; si esto no es posible, todavía pueden adquirir *algún* significado al ser utilizados en enunciados a partir de los cuales puedan derivarse otros enunciados que utilicen términos empíricos. De este modo, los términos teóricos no definidos absorberán, por así decirlo, algún significado empírico. El significado empírico se filtrará de los enunciados básicos y los términos básicos ascendiendo hasta los enunciados teóricos y los términos teóricos.

Si esta concepción fuera correcta, entonces nada podría tener mayor influencia sobre el significado de un término teórico que un cambio en su posición respecto a la contrastabili-

dad —o, más exactamente, en la posición respecto a la contrastabilidad de algunos enunciados en los que él aparezca— de «incontrastable» a «contrastable» y de ahí a «contrastado». Sin embargo, este cambio no afecta, en general, al significado, como hemos visto; y así se desmorona la teoría inductivista de la absorción de significado empírico por los términos teóricos.

Esto no tiene por qué sorprendernos, considerando que la distinción entre términos empíricos y términos teóricos es completamente errónea. Como he dicho en otras ocasiones, todos los términos son teóricos, aunque algunos son más teóricos que otros[15].

25. LOS TERMINOS METAFISICOS PUEDEN DEFINIRSE POR MEDIO DE TERMINOS EMPIRICOS

Pueden merecer la pena esbozar brevemente un procedimiento general por medio del cual podamos definir, en términos empíricos, los predicados metafísicos; es decir, predicados que sean esencialmente incontrastables (aunque quizá sean verificables).

Sea x una persona (o un animal o planta u otra cosa). Queremos definir una disposición metafísica de x, una habilidad de x que no sólo está «oculta» en el sentido de Berkeley, sino que, en principio, no es contrastable.

Tomemos cualquier cosa empírica y. Puede ser una cosa de lo más corriente. Sea un vaso de agua.

¿Qué hará x con un vaso de agua? Lo más sencillo es bebérselo. Así llegamos al esquema «x tiene la disposición D si y sólo si, *existe* un y tal que y sea un vaso de agua y si x bebe y, entonces...». Ahora insertemos en los puntos suspensivos algo emocionante, algo como rejuvenecerse o convertirse en león. Esto termina la historia: «x es capaz de rejuvenecerse (o convertirse en un león) si, y sólo si, existe un y tal que y sea un vaso de agua (muy especial; o una jarra de oro, llena de agua) y si x bebe y, x tendrá la apariencia de cuando era joven y se sentirá como entonces (o tendrá la apariencia de un león y sentirá como uno) y x conoce el secreto de y».

Si x no es una persona o un animal, sino una cosa, entonces podemos continuar trabajando con nuestro vaso de agua: x puede convertirse entonces en oro, o en una perla, si se moja con el agua o si se introduce en ella. Además, la disposición metafísica puede ser tal que la posean todos los hombres (o

[15] *L. I. C.*, final de la sección 25; véase también *El desarrollo del conocimiento científico (C. & R.)*, pp. 118, 388.

animales o plantas u otras cosas) o solamente algunos muy especiales (personas nacidas en domingo, o bajo el signo de Géminis; o una hierba plantada durante la luna nueva).

Por ejemplo, sea *x* una cierta enfermedad. Entonces *y* puede ser la hierba que la cura, si *y* se toma tres veces al día. Así llegamos a la propiedad metafísica «curable por *alguna* hierba que existe en *algún lugar*».

Pueden decirme que lo que he descrito no son predicados metafísicos, sino más bien mágicos. Sin embargo, la metafísica primitiva está estrechamente relacionada con la magia. Consiste en fórmulas mágicas protegidas contra la refutación *por su carácter existencial.*

Los predicados metafísicos cuya construcción se describe aquí tienen todos significado, porque pueden ser, todos, *definidos* mediante términos con significado, y generalmente se supone que una definición da significado a los términos definidos siempre que los términos definitorios tengan significado. Y, puesto que podemos escoger los términos empíricos más triviales como términos definitorios, está claro que no hay problema en este aspecto.

Pueden idearse definiciones más complejas, haciendo uso de más de un operador existencial: puede decirse que dos hierbas tienen la propiedad disposicional de convertirse en un curalotodo si, y sólo si, existe una constelación *y* si existe una fórmula latina secreta tales que si bajo esa constelación se cuecen juntas las hierbas mientras se pronuncia la fórmula, el resultado será un curalotodo.

El principal ataque de los positivistas es, desde luego, contra la metafísica religiosa, sobre todo contra la posibilidad de una «teología racional». Afirman que sus términos tienen que carecer de sentido. Para refutar esta aserción intenté, hace algún tiempo, formular lo que llamé «la aserción archi-metafísica» en el simbolismo de *Testability and Meaning* de Carnap[16]. Por «la aserción archi-metafísica» entendía yo la aserción de la existencia de un Dios personal, esto es, un espíritu omnipotente, omnipresente, omnisciente y personal. «*x* es omnipotente» puede definirse generalizado la idea «*x* puede poner la cosa *y* en el lugar *z*»: esta idea pertenece a la vida corriente y, por tanto, es suficientemente empírica. Y aunque su generalización, «*x* puede poner cualquier cosa en cualquier lugar» quizá no sea ya empírica, es definible mediante la aplicación de una operación lógica muy corriente a un término

[16] Esto lo intento en mi artículo «La demarcación entre la ciencia y la metafísica», en *El desarrollo del conocimiento científico: Conjeturas y refutaciones (C. & R.)*, capítulo 11. Véanse las notas 17 a 19 en la próxima sección.

empírico y, por tanto, está claro que tiene sentido. La omnipresencia es aún más fácil de definir mediante la generalización del predicado empírico «*x* está situado en el lugar *y*». El resultado es «*x* está situado en todas partes». Los términos «omnisciencia» y «persona» pueden definirse juntos, con ayuda del predicado «*x* conoce *y*». Podemos decir que *x* es una persona si, y sólo si, existen un *y* y *x* conoce *y*: y podemos decir que *x* es omnisciente si, y sólo si, *x* está situado en algún lugar y si *x* conoce todo *y*. Respecto al espíritu, puede decirse que una persona omnipresente es un espíritu. De este modo, podemos dar «significado empírico» a la teología racional. Sin embargo, sigue siendo metafísica: no porque carezca de sentido, sino porque sus aserciones no pueden ser contrastadas, porque son irrefutables.

Algunos positivistas han aceptado la significatividad de mi construcción del enunciado al que he llamado «la aserción archi-metafísica»; pero han replicado que mi enunciado, como tiene significado, está claro que no es metafísico, sino sencillamente, un *enunciado empírico falso;* a lo que añaden el aditamento de que éste es un viejo problema resuelto hace ya mucho tiempo. El aditamento es una parte necesaria de la réplica, y es falso. Porque puede mostrarse fácilmente, con la literatura del Círculo de Viena, que el principal ataque del positivismo fue sobre la «metafísica tradicional» de la cual, la llamada «teología racional», como la de Spinoza, era el ejemplo más destacado. La tesis positivista era que todo esto *carecía irremediablemente de sentido.* Esta tesis la descartan ahora, claramente, los que dicen que, al mostrar que es posible interpretar algunas teorías metafísicas como dotadas de sentido, no he hecho más que mostrar que no son metafísicas.

Entonces, podemos preguntar nosotros ahora: «*¿Qué era esa filosofía tradicional que ustedes atacaron un día tan violentamente?* ¿A qué venía tanto acaloramiento? Y, en cualquier caso, ¿no relacionaron ustedes la contrastabilidad con el sentido? Y nuestros ejemplos, y otros, ¿no muestran que existen enunciados incontrastables pero con sentido? ¿Y por qué la tautología 'una combinación de palabras sin sentido es una combinación de palabras sin sentido' había de tener el más mínimo interés para nadie bajo el Sol? Y sin embargo, la aserción de ustedes de que la metafísica carece de sentido, ¿no se convierte en esa tautología, si responden a mis ejemplos diciendo que, puesto que tienen sentido, no son metafísicos?».

Pero, dejando a un lado las polémicas, puedo resumir mi concepción como sigue: mientras que la mayor parte de los *conceptos científicos* no pueden ser definidos en términos de lo que pueden llamarse «fenómenos», ni siquiera en lo que puede

llamarse los términos empíricos del lenguaje corriente, algunos *conceptos metafísicos* sí que pueden definirse de ese modo. En consecuencia, un intento de caracterizar a la ciencia (en contraposición con la metafísica) por un criterio como la definibilidad empírica de sus términos lleva a una demarcación que es, al mismo tiempo, demasiado estrecha y demasiado amplia: excluirá casi todo lo que se quiere que incluya e incluirá mucho de lo que se quiere excluir.

26. LA CAMBIANTE FILOSOFIA DEL SENTIDO Y EL SIN SENTIDO

Wittgenstein, en su *Tractatus*, como Berkeley, tenía una filosofía del significado y de la carencia de significado, del sentido y del sin sentido, que era vigorosa y clara. De un lado se encontraban los enunciados empíricos informativos y del otro un mero galimatías, mera verborrea; aunque una verborrea que podía tener la apariencia de un enunciado empírico. Los signos que constituían los enunciados empíricos eran palabras a las que se había dado un significado empírico: cada una de ellas estaba asociada, por el uso corriente, con ciertas cosas o sucesos observables. El galimatías, por otro lado, o bien era antigramatical (como en el caso de «Sócrates es idéntico») o contenía palabras a las que «no habíamos dado un significado». Y la sola tarea de la filosofía era —según Wittgenstein— «demostrar» a la gente que hablaba de metafísica que estaba hablando sin sentido. Así es como se resumió su mensaje en el *Tractatus* (6.53): «El método correcto de la filosofía sería éste: no decir nada, excepto lo que puede decirse, esto es, las proposiciones de la ciencia natural, esto es, algo que no tiene nada que ver con la filosofía: y luego, siempre que alguien quisiera decir algo metafísico, demostrarle que no había dado significado a ciertos signos de sus proposiciones».

Aquí había una teoría provocadora. Era claro que estaba incompleta mientras no hubiera mostrado que las palabras o signos o conceptos que se usan en las ciencias naturales pueden, todos ellos, definirse por medio de definiciones empíricas, definiciones basadas en conceptos que expresen experiencias inmediatas. Esta fue la tarea que se impuso Carnap a sí mismo en *Der Logische Aufbau der Welt* (1918). En este libro trató de desarrollar los esquemas de un sistema de definiciones o «constituciones» de los términos científicos: un concepto estaba «constituido» si existía una *cadena de definiciones* que lo redujeran a términos de la experiencia inmediata.

Esta era, naturalmente, una teoría de la inducción, a partir

de «datos», aplicada a los conceptos en vez de serlo a los enunciados. Pero se quedó en programa, porque la cosa no podía hacerse. La «constitución de conceptos» no era más posible que la «inducción de teorías»[17].

Como consecuencia de ello, la sencilla y directa teoría de Wittgenstein sobre la carencia de significado o el sin. sentido tuvo que ser abandonada. Se abandonó; y la filosofía de la carencia de significado o del sin sentido y, con ella, la escuela positivista empezaron a desintegrarse.

A lo largo de los años se han intentado muchos pequeños cambios y ajustes. La directa idea de la falta de significado o de galimatías ha sido sustituida sucesivamente por otras varias ideas más sofisticadas[18]. Se ha dicho, por ejemplo, que una expresión es «metafísica» si «no es una cláusula significativa del lenguaje de la ciencia» (como si a algún metafísico fuese a importarle si usaba o no «el lenguaje de la ciencia», o lo que alguien pretendía que era tal lenguaje). Así, la tesis de la falta de significado se ha ido aguando hasta quedar irreconocible, excepto para los expertos. Una cosa similar ha ocurrido con la idea de metafísica. Expresiones que durante la década de 1930 se hubieran considerado modelos de sin sentido metafísico (ya que no eran ni verificables ni falsables), se consideran ahora perfectamente respetables. Y todos esos cambios subrepticios han sido introducidos *ad hoc* para evitar las consecuencias no buscadas de un criterio de significado que, por mucho que se haya modificado, ha seguido siendo inadecuado. Porque fue siempre *demasiado estrecho y demasiado amplio* al mismo tiempo, como puede mostrarse fácilmente: siempre ha situado las teorías científicas y las teorías metafísicas indeseables a los lados equivocados de la línea de demarcación. Porque incluso la teología racional, por ejemplo, puede «formalizarse» en alguno de los lenguajes de la ciencia propuestos[19].

---

[17] Señalé esto en mi *L. I. C.* (final de la sección 25) y lo había señalado antes aún en mis discusiones con los miembros del Círculo de Viena. Véanse por ejemplo la nota 27 y el texto entre las notas 25 y 27 de mi artículo «La demarcación entre la ciencia y la metafísica», *op. cit.*

[18] Cuento la historia con más detalle en mi artículo «La demarcación entre la ciencia y la metafísica», *op. cit.* [*Cf. Búsqueda sin término*, secciones 16 y 17. *Ed.*]

[19] Véase la nota 16 *supra*, y también mi artículo «La demarcación entre la ciencia y la metafísica», *op. cit.* Carnap, desde entonces, ha escrito un artículo «The Methodological Character of Theoretical Concepts» (véanse mis referencias en las notas 93 y 94 a la sección 11 y a la nota 112 en la sección 13), en el que revisa algunas de las concepciones que critiqué en mi artículo; sobre todo, trata de dar un nuevo criterio de significado (ahora llamado «criterio de significación») y ofrece dos pruebas, una que establece que su criterio *no es demasiado amplio* (pp. 54-57) y la segunda, que *no es demasiado estrecho* (pp. 58-59). Con todo, en presencia de ciertas suposiciones que se hacen en la segunda

¿Por qué no abandonar simplemente el dogma según el cual el criterio de significado (o —ahora— de «significación») debe ser equivalente al criterio de demarcación? Los intentos de salvar este dogma han causado interminables molestias. No puedo imaginar por qué se consideró que merecía la pena tomarse la molestia. ¿Quizá para salvar a la venerable filosofía del positivismo de ser refutada? Pero ha sido refutada; y ha sido abandonada, de una u otra forma, por todos los que entienden la situación. Así que ¿por qué seguir molestándose?

Con esto concluyo mis comentarios sobre el problema de la demarcación, puesto que la mayoría de sus aspectos lógicos, especialmente la conexión con la idea de *contenido* empírico, se han tratado ya bastante satisfactoriamente en *L. I. C.* y no están urgentemente necesitados de mayor dilucidación.

---

prueba, puede mostrarse que la primera prueba se desbarata y que el criterio resulta *demasiado amplio;* y, del mismo modo, en presencia de ciertas suposiciones que se hacen en la segunda prueba, se desbarata la segunda, y el criterio resulta *demasiado estrecho.*

CAPITULO IV

LA CORROBORACION

27. CORROBORACION: CERTIDUMBRE, INCERTIDUMBRE, PROBABILIDAD

En los dos capítulos anteriores he explorado las ramificaciones lógicas del problema de la inducción. Hay otra ramificación, sin embargo, que todavía no he tocado, porque no está *lógicamente* conectada con el problema de la inducción. Pero está relacionada con él por medio de lazos que pueden resultar más fuertes incluso que los de la lógica: por el prejuicio inductivo y *por una solución errónea del problema de la inducción* que, desgraciadamente, sigue siendo ampliamente aceptada como válida. Estoy aludiendo, naturalmente, a la concepción según la cual, aunque la inducción es incapaz de establecer una hipótesis inducida con *certeza,* es capaz de hacer algo casi tan bueno como eso: puede atribuir a la hipótesis inducida cierto grado de *probabilidad* (y una probabilidad 1 sería certeza).

Esta concepción es un error radical; no obstante, puede apoyarse en un argumento muy convincente. Este argumento puede presentarse como sigue.

Todo el problema de la inducción, según ese argumento, surge claramente del hecho de que las inferencias inductivas *no son válidas:* lo que equivale a decir que las conclusiones inductivas no se derivan deductivamente de premisas inductivas. Pero no hay por qué alarmarse por este hecho algo trivial; sobre todo en vista de que existe una clase amplia e importante de inferencias en las que la conclusión no se deriva estrictamente de las premisas. De hecho, cada inferencia deductiva puede modificarse para que dé lugar a una inferencia que no sea válida, sino solamente «más o menos válida», o «válida en cierto grado». Tomemos el ejemplo siguiente:

VALIDA	VALIDA EN CIERTO GRADO
Todos los hombres fuman	*x* por ciento de todos los hombres fuma
Jack es un hombre	Jack es un hombre
Jack fuma.	Jack fuma.

Como sugiere el ejemplo, podemos decir que, mientras que en la inferencia *válida,* la verdad de las premisas hace *cierta* la verdad de la conclusión, en la inferencia *más o menos válida,* la verdad de las premisas deja la verdad de la conclusión *incierta.* Sin embargo, la conclusión es más o menos *probable* —para ser precisos, probable en el grado de *x* por ciento—.

Las inferencias que establecen una conclusión como probable —o las inferencias probables, como podríamos llamarlas— existen. Como lo muestra el ejemplo, son inferencias en las que la conclusión, de uno y otro modo, va más allá de lo que enuncian las premisas: no está plenamente entrañada por ellas. Evidentemente, esto ocurre con todas las conclusiones inductivas. Esto parece mostrar que se resolverá el problema de la inducción tan pronto como hayamos desarrollado una teoría de la probabilidad que nos permita evaluar la probabilidad de una conclusión inductiva —es decir, de una hipótesis— dadas ciertas premisas inductivas o evidenciales. Si escribimos

$$p\ (a,b) = r$$

con el significado de «la probabilidad de *a* dado *b* es igual a *r*» (donde *r* es alguna fracción entre 0 y 1), entonces podemos dar al argumento convincente la forma que sigue:

Sea *h* una hipótesis (una «conclusión inductiva») y *e* nuestra «evidencia inductiva». Entonces el problema de la inducción consiste —o eso parece— en determinar el valor de *r* en

$$p\ (h,e) = r$$

esto es, el valor de la probabilidad de la hipótesis inductiva *h* dada la evidencia *e*. Así, el problema de la inducción se resolverá construyendo una lógica generalizada, una lógica de la probabilidad. *Porque según este persuasivo argumento, la lógica inductiva no es nada más que lógica probabilística.* Es la lógica de la inferencia incierta, del conocimiento incierto; y *p (h,e)* es el grado en el que nuestro conocimiento cierto [seguro] de la evidencia *e* justifica racionalmente nuestra creencia en la hipótesis *h*[1].

---

[1] En *L. I. C.* (Apéndice *IV) di un número de sistemas de axiomas para el cálculo formal de probabilidades (una de cuyas interpretaciones es la interpretación lógica). Aunque no supongo ningún conocimiento por parte del lector de tales tecnicismos, enunciaré aquí, para completar mi exposición, uno de los sistemas de axiomas más sencillos. Consta sólo de tres axiomas; el primero introduce la idea *p(a,b)*, el segundo la idea del producto *ab* (que debe leerse «*a*-y-*b*»); y el tercero la idea del complemento $\bar{a}$ (que debe leerse como «no-*a*»).

A existen al menos dos elementos *a* y *b*, tales que

$$p(a,a) \neq p(a,b)$$

Como ya he dicho antes, creo que este argumento es totalmente erróneo. La apelación a la probabilidad no afecta en absoluto al problema de la inducción. (*Cf.* las secciones 1, 80 y 83 de *L. I. C.* y los pasajes de Hume en el Apéndice *VII, texto de las notas 4 a 6.) Formalmente, puede sustentarse esto con la observación de que toda hipótesis universal h va tanto más allá que cualquier evidencia empírica e, que su probabilidad $p\ (h,e)$ siempre será cero, porque la hipótesis universal hace afirmaciones sobre un número infinito de casos, mientras que el número de casos observados sólo puede ser finito.

Hay como mínimo dos formas de criticar la supuesta conexión lógica entre los problemas de la inducción y de la probabilidad. Dejo la discusión completa de uno de ellos para la

---

B $$p(ab,c) = p(a,d)p(b,c) \leq p(a,c)$$

siempre que $p(a,a) = p(bc,d)$ y $p(bc,c) = p(d,c)$.

C $$p(\bar{a},b) = p(c,c) - p(a,b),$$

siempre que $p(d,d) \neq p(d,b)$ para algún elemento d.

A esto podemos añadir un postulado que introduzca la identidad (de sustitución) $a = b$, de dos elementos a y b:

(*) $a = b$, y sólo si, $p(a,c) = p(b,c)$ para todo elemento c. (En ese caso, a y b pueden sustituirse en todas parte el uno por el otro.)

Pueden añadirse nuevas definiciones adoptando una regla de definición basada en $a = b$, por ejemplo, podemos definir «a v b» (léase «a —o— b»), como es usual, por medio de

$$\text{«}a \text{ v } b = \overline{\bar{a}\bar{b}}\text{»}.$$

Este sencillo sistema de axiomas nos permite deducir, además de todas las leyes del álgebra booleana, las bien conocidas leyes del cálculo de probabilidades, tales como la *ley de reflexividad*.

1) $$p(a,a) = p(b,b) = 1$$

Consecuencia inmediata de 1) y B son la muy importante *ley de monotonía*

2) $$p(ab,c) \leq p(a,c)$$

y la *ley general de la multiplicación*,

3) $$p(ab,c) = p(a,bc)p(b,c)$$

Una consecuencia más remota del sistema es la *ley general de la suma*,

4) $$p(a \text{ v } b,c) + p(ab,c) = p(a,c) = p(b,c)$$

Obtenemos, además, una serie de teoremas que no pueden obtenerse en el cálculo habitual, por ejemplo, los siguientes:

5) $$p(a,\overline{\bar{b}\bar{b}}) = 1$$

6) $ab = b$ si, y sólo si, $p(a,b\bar{a}) \neq 0$.

Dentro de la interpretación lógica, la igualdad «$a = b$» puede leerse «es lógicamente equivalente a» (o es «inter-deducible con») y $ab = b$ puede entonces leerse «a se sigue de b» lo que puede escribirse, naturalmente, «$a \geq b$» o «$b \leq a$».

segunda parte de este tomo, donde me propongo criticar *la interpretación subjetiva de la probabilidad*, es decir, la interpretación de la probabilidad como grado de conocimiento incompleto o incierto. Trataré de mostrar allí que la idea de probabilidad inductiva no funciona y por qué no puede funcionar.

La otra forma consiste en señalar que *la creencia intuitiva en la probabilidad inductiva de una hipótesis a la luz de la evidencia* es una mezcla de, al menos, dos ideas intuitivas. Una de esas ideas —que se discutirán en ésta y en las secciones siguientes— es defendible, pero sólo tiene una importancia limitada y no está lógicamente conectada con el problema de la inducción; mientras que la otra —la idea de una *lógica de la probabilidad inductiva* que acabo de mencionar— es indefendible. Que es indefendible por razones lógicas y matemáticas se mostrará con más detalle en la segunda parte de este volumen; aquí quiero señalar otro aspecto de esta idea, uno que me parece a mí en cierto modo más importante. Me refiero a la equivocada actitud hacia la ciencia de la que parece surgir esta idea.

Ocupémonos en primer lugar de la idea que creo defendible. Es la idea de que las hipótesis pueden diferenciarse según los resultados de sus contrastaciones: la idea de que algunas hipótesis están bien contrastadas por la experiencia y otras no tanto; que hay, además, hipótesis que, por el momento, no han sido contrastadas en absoluto, e hipótesis que no han resistido las contrastaciones y que, por tanto, pueden considerarse falsadas. Si consideramos una serie de hipótesis desde este punto de vista, no puede oponerse objeción alguna a su clasificación según el grado en el que han resistido sus contrastaciones —igual que podemos dar notas a los alumnos que han hecho una serie de exámenes, unos fáciles y otros difíciles—.

El deseo de calificar a las hipótesis según las contrastaciones que hayan pasado con éxito es legítimo: no conozco ninguna objeción seria. Por razones que expondré en la próxima sección, propongo llamar a la «nota» de una hipótesis, o el grado en el cual ha resistido a las contrastaciones, su «*grado de corroboración*» (más que su «probabilidad»). Uno podría pensar que es muy importante, ya que *la aceptabilidad de una hipótesis* dependerá, obviamente, de su grado de corroboración. Pero no creo que ninguna inquietud ni diferencia de opinión respecto a la aceptabilidad de una hipótesis vaya a ser nunca eliminada por medio de una determinación «exacta» de su grado de corroboración. Aunque daré más adelante una definición del grado de corroboración —una que nos permita comparar teorías rivales tales como la de Newton y la de Einstein— dudo que una evaluación numérica de este grado tenga ninguna significación práctica. (Quizá no sea yo objetivo, porque, en gene-

ral, no tengo mucha fe en las notas de los exámenes.) Sea como fuere, la idea de un grado de corroboración es significativa, en nuestro contexto actual, principalmente a causa de la ayuda que puede prestar para aclarar el gran embrollo creado por concepciones equivocadas sobre las probabilidades inductivas.

Esto me trae a mi principal y último punto en esta sección: el transfondo filosófico general de la creencia en las probabilidades inductivas.

La concepción sobre la ciencia en la que se origina esta creencia es fundamentalmente la vieja concepción de la Ciencia, con mayúscula. Es la concepción de la ciencia como *scientia* o *epistèmè*, un conocimiento cierto, demostrable. Sin duda alguna, esta concepción se ha modificado un poco ahora: todo el mundo se da cuenta ahora de que una certeza plena es inalcanzable en las ciencias que se llaman «inductivas». Pero como la inducción se considera un tipo de (debilitada) generalización de la deducción, el antiguo ideal se ha modificado sólo ligeramente.

Este antiguo ideal, sin embargo, debe abandonarse por completo. Debería abandornarse incluso aunque sólo pensásemos en los sistemas puramente deductivos. Ya no consideramos sistema deductivo a uno que establece la verdad de sus teoremas deduciéndolos de «axiomas» cuya verdad es totalmente cierta (o evidente por sí misma o fuera de duda); más bien consideramos deductivo a *un sistema que nos permite argüir sus diversos supuestos racional y críticamente,* averiguando sistemáticamente sus consecuencias. La deducción no se usa simplemente con el propósito de *demostrar* unas conclusiones; se utiliza más bien como instrumento de la crítica racional. Dentro de una teoría puramente matemática, deducimos conclusiones para investigar el poder y la fertilidad de nuestros axiomas. Y dentro de una teoría física, deducimos conclusiones para criticar, y sobre todo, contrastar, las conclusiones deducidas y con ello, nuestras hipótesis: por regla general, no tenemos intención de *establecer* nuestras conclusiones.

Al considerar la probabilidad inductiva como una medida de la sensatez de nuestras creencias o de la fiabilidad de nuestro conocimiento, el devoto de la inducción probable deja claro que todavía sigue aferrado, como Bacon, a un ideal debilitado de *epistèmè*. Concibe sus enunciados evidenciales *e* como dotados de un papel análogo al de los axiomas evidentes por sí mismos que se suponía que «demostraban» nuestros teoremas. Y concibe su hipótesis *h* como dotada de un papel análogo al de los teoremas cuya verdad se hace cierta por deducción a partir de los axiomas; sólo que, al ser más débil la inducción que la deducción, ahora no obtenemos más que un sucedáneo

de certeza: la probabilidad entra como sustituto de la certeza: no es del todo lo que buscábamos, pero es lo más parecido a ello y, en todo caso, algo cercano.

Todo esto es inaceptable. Los propios enunciados evidenciales *e* están lejos de ser ciertos. [Siempre he puesto este punto de relieve; no obstante, los inductivistas lo usaron como argumento contra mi doctrina sobre la asimetría entre la falsación y la verificación; *cf.* sección 22. Pero ningún inductivista ha explicado nunca cómo interpretar «*p (h,e)*» cuando *e* mismo es incierto y, presumiblemente, «sólo probable»][2]. Tampoco nos son «dados» esos enunciados evidenciales por Dios, por la naturaleza, por nuestros sentidos. Toda observación y, aún en mayor grado, todo enunciado de observación es, él mismo ya, *una interpretación a la luz de nuestras teorías*[3]. Sin embargo, a pesar de que este hecho tiene gran importancia, plantea una cuestión menor en comparación con lo que quiero criticar aquí: una actitud general; una filosofía general de la ciencia; una filosofía que considera que su principal problema es el de explicar de dónde deriva la ciencia su «certeza», su fiabilidad racional, su validez o su autoridad. Porque yo sostengo que la ciencia no posee certeza, ni fiabilidad racional ni validez ni autoridad. Lo mejor que podemos decir de ella es que, aunque consisten en nuestras suposiciones, nuestras conjeturas, estamos haciendo todo lo que podemos por contrastarlas; es decir, por criticarlas y refutarlas.

Pero la filosofía inductivista no sólo atribuye autoridad a la ciencia, sino también (quizá sin proponérselo) atribuye a la ciencia un planteamiento cauteloso, tímido incluso, que es totalmente ajeno a nuestro procedimiento real. Esta filosofía, al considerar como objetivo de la ciencia el alcanzar altas probabilidades para sus teorías, supone que la ciencia actúa según la regla: «¡Ve lo menos lejos posible de tu evidencia *e*!». Porque el contenido de la hipótesis *n* no puede ir muy lejos de la evidencia *e* sin reducir *p (h,e)* a un valor muy próximo a cero. Por ejemplo, sea *e* la conjunción de muchas descripciones de un suceso de cierto tipo: *n* no necesita afirmar muchos otros sucesos, no incluidos en *e*, para hacer *p (h,e)* muy pequeña, según el cálculo de probabilidades. Esto muestra que una probabilidad elevada es la dudosa recompensa por decir muy poco o nada. En otras palabras, la regla «¡Obtén probabilidades altas!» supone apreciar las hipótesis *ad hoc*.

Todo esto ofrece una imagen de la ciencia poco inspiradora,

[2] [Véase R. C. Jeffrey, *The logic of Decision*, 1965, capítulo 11. *Ed.*]

[3] [*Cf.* Donald T. Campbell, «Evolutionary Epistemology», en P. A. Schilpp, ed., *The Philosophy of Karl Popper*, pp. 413-463. *Ed.*]

una imagen, además, que no se parece al original. En realidad, lo que hace que el original sea tan inspirador es su audacia; sus hipótesis audazmente concebidas, audazmente sometidas a todo tipo de crítica —a todas las refutaciones que podemos idear, incluidas las contrastaciones mas severas que nuestra imaginación nos ayuda a concebir—. Es esta audacia la que nos permite transcender los límites, estrechos al principio, de nuestra imaginación y del lenguaje corriente.

28. ¿«CORROBORACION» O «PROBABILIDAD»?

En la sección anterior introduje el término «grado de corroboración» para caracterizar el grado en el que una hipótesis ha resistido a las contrastaciones. En esta sección me propongo discutir una mera cuestión terminológica: mis razones para proponer que se hable de «grado de corroboración» en vez de hablar de la «probabilidad de una hipótesis a la luz de las contrastaciones». Mi principal razón es, naturalmente, que esta última frase —aunque perfectamente legítima en sí— puede inducir a confusión.

Estoy dispuesto a admitir que las palabras «probable» y «probabilidad» se usan con frecuencia en un sentido similar al sentido en que propongo usar la expresión, algo desmañada, de «grado de corroboración». Una expresión tal como «una suposición probable» pertenece al habla corriente; y expresiones como «una conjetura improbable» o quizá incluso «una hipótesis improbable» son expresiones perfectamente directas del uso normal, aunque quizá ligeramente «erudito». Además, el sentido ordinario de una frase como «debe usted tratar de pensar en una hipótesis más probable» o «de las diversas hipótesis que se han presentado hasta el momento, la suya me parece, a la luz de las contrastaciones, la más probable», la encuentro suficientemente clara: el sentido que se le da está estrechamente relacionado, sin duda, con el que propongo para el término más artificial «grado de corroboración»; una respuesta que nos haría cambiar la última frase por ésta, más desmañada: «de las diversas hipótesis presentadas hasta el momento, la suya me parece ser la que, a la luz de las contrastaciones, tiene el más alto grado de corroboración». No obstante, en *L. I. C.* (secciones 79 a 84) descarté el uso establecido de las palabras «probable» y «probabilidad» y los sustituí por una terminología más artificial. Esta sustitución requiere una defensa explícita, porque creo firmemente que debemos hablar de una forma sencilla y clara, y no debemos desviarnos del uso ordinario sin razones poderosas.

Hay razones poderosas, sin embargo, para hacerlo aquí; y en vista de que sólo se dan de forma implícita en *L. I. C.* (secciones 81 a 83) trataré de enunciarlas aquí con todo detalle.

1) Hay un segundo uso, mejor establecido aún, de «probable» y «probabilidad» que apenas se distingue, intuitivamente, del primero. (Con el primer uso me refiero a esa «probabilidad» de las hipótesis a la que propongo que se re-bautice como «grado de corroboración».) Aparece en frases como: «probablemente Ana se casará con Arturo» o «es probable que Bob pase los exámenes finales por los pelos». Esta es la probabilidad que puede describirse como la *probabilidad de un suceso*. Tiene un rasgo característico: la probabilidad de que Ana se case con Arturo y de que Bob pase los finales por los pelos no puede exceder la probabilidad del menos probable de los dos sucesos, y normalmente es menor que ella. De forma más general, *la probabilidad de un suceso complejo, consistente en la concurrencia de varios sucesos singulares, en general, será menor o, como mucho, igual a la probabilidad de cualquiera de los sucesos componentes.*

Para los enunciados (hipótesis) que describen sucesos, lo siguiente es válido, en consecuencia: la probabilidad de un enunciado que describe un suceso *disminuye al aumentar el contenido lógico del enunciado.*

Esta «norma de contenido», como puedo llamarla (corresponde al «axioma de monotonía» del cálculo de probabilidades; véase la sección anterior, nota 1 y *L. I. C.*, Apéndices *IV y *V), concuerda con el uso lingüístico establecido, según el cual, una tirada de 12 con dos dados se considera «menos probable» que una tirada de 6 con un dado.

No hay una regla de contenido que corresponda al primer sentido de «probabilidad», al que yo he llamado «grado de corroboración». Al contrario, la mayoría de los físicos dirán que la teoría de Maxwell sobre la luz es «más probable» en el primer sentido —es decir, está «mejor corroborada» o «mejor contrastada»— que la teoría de la luz de Fresnel. La razón es que la teoría de Maxwell ha sido *contrastada más extensa y rigurosamente,* incluso en campos en los que no puede contrastarse la de Fresnel. Al mismo tiempo, la teoría de Maxwell tiene un contenido lógico mucho mayor que la de Fresnel: la de Maxwell es una teoría ondulatoria de la luz y una teoría del electromagnetismo, mientras que la de Fresnel no es más que una teoría ondulatoria de la luz. Así pues, la teoría de Maxwell, aunque es «más probable» en el sentido de que está «mejor corroborada» o «mejor contrastada» es, al mismo tiempo, «menos probable» en el sentido del segundo uso de la palabra, que también está bien establecido, sobre todo si no pensamos

tanto en *las contrastaciones superadas con éxito por una hipótesis* como en las *oportunidades de que ocurra un suceso.*

2) Los dos usos —la probabilidad de una hipótesis en relación con sus contrastaciones y la probabilidad de un suceso (o una hipótesis) en relación con sus oportunidades para ocurrir— se ha diferenciado raras veces y en general se tratan como iguales. Esto puede deberse a que intuitivamente —es decir, al menos «a primera vista»— son difíciles de distinguir. Pero puede mostrarse que mientras que *el segundo uso concuerda con las reglas del cálculo matemático de probabilidades (especialmente con su «axioma de monotonía»)*, el primero no concuerda.

Un ejemplo temprano, y especialmente instuctivo, de no diferenciación entre los dos sentidos se encuentra en una famosa carta de Leibniz[4]. En esta carta, Leibniz habla muy claramente de hipótesis a las que yo describiría diciendo que tienen mucho *contenido*; habla de su «simplicidad», su «virtud» y su «poder» (explicativo); y dice que una hipótesis es más «probable» cuanto mayores son su simplicidad y su poder, y cuanto mayor es el número de fenómenos que puede «resolver» (explicar) con el menor número posible de «suposiciones» (adicionales o *ad hoc*). Está claro que tiene en mente nuestro primer sentido de probabilidad, lo que propongo que llamemos «grado de corroboración». Sin embargo, parece pensar que esta «probabilidad» se comporta como la que satisface el cálculo de probabilidades, porque sugiere que esta probabilidad es una especie de sustituto o *Ersatz* de la certidumbre o de la verdad, puesto que puede aproximarse a la «certeza física» (o «certeza moral», como dice Couturat en sus comentarios). No hay ninguna sugerencia de que Leibniz perciba que la simplicidad y el poder explicativo son equivalentes a la *improbabilidad* lógica, en el sentido del cálculo de probabilidades (*cf. L. I. C.*, secciones 34 a 46, 83, 30 y 32).

3) Ahora bien, si estoy en lo cierto en todo esto, será importante evitar una confusión entre estos dos sentidos: sobre todo considerando que casi todo el que ha escrito sobre la probabilidad de las hipótesis ha supuesto, *sin mayor discusión,* que la probabilidad de las hipótesis con respecto a las *contrastaciones* superadas por ella (o su «grado de corroboración») podía tratarse en términos del cálculo de probabilidades.

Si, por el contrario, estoy equivocado al afirmar que el grado de corroboración no satisface el cálculo de probabilida-

[4] Carta a Conring, 19 de marzo de 1678; véase vol. I, pp. 195 s. de la edición de Gerhart de *Die philosophischen Schriften von G. W. Leibniz*, 1875-1890, y Louis Couturat, *La Logique de Leibniz*, 1901, p. 268.

des, entonces tampoco hace ningún daño distinguir provisionalmente entre los dos usos tan exactamente como sea posible, quizá para demostrar más tarde que ambos pueden tratarse en términos del cálculo de probabilidades y establecer de ese modo que yo estaba equivocado. (Creo, sin embargo, que puedo *demostrar* que tengo razón; véase más adelante, sección 31.)

Así pues, todo habla en favor de una distinción clara, al menos provisionalmente, y en contra de una suposición prematura y acrítica de que el grado de corroboración debe satisfacer las reglas generales del cálculo de probabilidades.

4) Una razón adicional surge del hecho de que hay un tercer sentido de «probabilidad» —la probabilidad de una inferencia, explicada brevemente en la sección anterior— del que puede considerarse (como el segundo sentido) que satisface las reglas del cálculo de probabilidades. Es este tercer sentido el que tan frecuentemente se ha confundido, inconsciente y acríticamente, con el grado de corroboración; lo confundieron, por ejemplo Locke y Hume, que distinguieron una «inferencia probable» o un «argumento probable» de un «argumento demostrativo» o una «prueba»[5].

Ahora bien, si tengo razón al decir que el grado de corroboración no satisface las reglas del cálculo de probabilidades, entonces las probabilidades de Locke y Hume tienen que ser estrictamente diferenciadas de la probabilidad en nuestro primer sentido, es decir, del grado de corroboración.

5) Si han de introducirse nuevos términos artificales para sustituir las palabras «probable» y «probabilidad» en uno u otro de los sentidos discutidos aquí, será mejor escoger un nuevo nombre para el primer sentido y hablar, por ejemplo, de una «hipótesis corroborada» por la siguiente razón:

La probabilidad de un suceso es el concepto más popular: está establecido en los juegos de azar, así como en las ciencias físicas y sociales y, desde luego, en matemáticas. La «probabilidad» de una ley universal a la luz de las contrastaciones es, en comparación, un concepto del lógico o del filósofo. Y es menos objetable (si lo es en absoluto) hacer cambios en el uso filosófico que en el uso común. (También, quizá sea *ligeramente más probable* que la tentativa tenga éxito, en caso de que lo tenga.)

6) Todos los usos mencionados, menos el primero, satisfacen las reglas del cálculo de probabilidades. Mi sugerencia, por tanto, tiene la ventaja de que conduce a la siguiente regla terminológica, fácil de recordar:

---

[5] En *Enquiry* (*cf.* sección IV, primera nota a pie de página), aunque menos estrictamente que en el *Treatise*, Hume divide los argumentos en «*demostraciones, pruebas y probabilidades*».

Retenemos los términos «probable» y «probabilidad» en aquellos casos en los que no hay razones para dudar de que el cálculo matemático de probabilidades sea satisfecho (y en especial, la «regla de contenido» o «axioma de monotonía»). Para el resto de los casos —tal como la «probabilidad» de las hipótesis en nuestro primer sentido— escogemos otros términos tales como «corroboración», al menos provisionalmente: es decir, hasta que la duda sobre si las reglas del cálculo de probabilidades son o no aplicables a ellos, haya sido eliminada.

Con esto concluye mi defensa de las decisiones terminológicas tácitamente adoptadas en las secciones 79 a 83 de *L. I. C.* Una lectura de esas secciones mostrará que las razones aducidas aquí no sólo están ya implícitas allí, sino que algunas de ellas adoptan la forma de una crítica detallada de autores que, como Keynes, identifican acríticamente el grado de corroboración y la probabilidad.

En las secciones restantes de este capítulo no me ocuparé más que del grado de corroboración de las hipótesis, como lo llamaré de ahora en adelante, y evitaré el término «probabilidad», de acuerdo con mi regla terminológica.

Las otras probabilidades, sobre todo la probabilidad de sucesos y la de inferencias (lógica de la probabilidad) se tratarán en la Parte II de este tomo. (Véase también *L. I. C.*, Apéndice *IV.)

29. ¿CORROBORACION O CONFIRMACION?

En la sección anterior expliqué brevemente las razones que tengo para sospechar que *el grado de corroboración no satisface el cálculo de probabilidades.*

Esta sospecha crea todo nuestro problema; y para no oscurecerlo con el uso de una terminología confusa, introduje el concepto de «grado de corroboración», y recomendé con fuerza que el uso del término «probabilidad» se dejase para conceptos que satisfacieran el cálculo de probabilidades (y, sobre todo, la «regla de contenido», que corresponde a los que he llamado «el axioma de monotonía» del cálculo.)

Desgraciadamente, existe el peligro de otra confusión terminológica. Hasta hace poco tiempo (de hecho, hasta que *L. I. C.* estuvo en galeradas), no utilicé el término «grado de corroboración» sino, en su lugar, el de «grado de confirmación». E hice uso de este término precisamente por la misma razón: debido a la necesidad de evitar el término «probabilidad». Por tanto, tengo que aclarar ahora por qué he decidido cambiar mi termi-

nología, después de haberla utilizado por lo menos en media docena de publicaciones.

La historia es bastante sencilla. Hasta hace poco, usé el término «grado de confirmación» porque era la traducción de Carnap, en su *Testability and Meaning* de mi término «*Grad der Bewährung*» o «*Bewährungsgrad*», el término que introduje en *L. d. F.* para evitar «probabilidad», por las razones que acabo de exponer[6]. Como las etiquetas no tienen importancia, no vi razones para no aceptar la traducción de Carnap, aunque no me gustaba demasiado. El término se estableció en seguida. Sin duda alguna, Carnap lo usó originariamente para denotar lo que yo quería que denotase: el grado en que una teoría ha resistido a las *contrastaciones*. (Incluso en un artículo reciente, Carnap sigue hablando del «requisito de confirmabilidad o contrastabilidad», con una referencia a *Testability and Meaning*, donde esos dos términos son traducciones de términos usados en *L. d. F.*, en el título de mi sección 83, por ejemplo.) Sin embargo, pronto empezó a usarse el término con un sentido nuevo y diferente; Carnap supuso, sin más discusión, en la primera oración de su libro *Logical Foundationes of Probability* que el «grado de confirmación» de una hipótesis satisfacía las reglas del cálculo de probabilidades.

Ninguno de los participantes en este acontecimiento mencionó jamás que yo había argüido extensamente en contra de ese punto y ninguno de mis argumentos fue replicado. Al parecer, habían caído en el olvido.

Encontré que ésta era una situación algo embarazosa y cuando publiqué un artículo (titulado «Grado de confirmación») en el que daba una definición de lo que *ahora* propongo llamar «*grado de corroboración*» me referí a esta ocurrencia en una nota a pie de página. En ella protesté levemente contra una observación de Carnap respecto al hecho de que era «generalmente aceptado» que el grado de confirmación satisfacía las reglas del cálculo de probabilidades.

Mi artículo obtuvo una réplica de J. Kemeny (que también usa «grado de confirmación» en el sentido de «probabilidad») en el *Journal of Symbolic Logic*. Decía, en el curso de su réplica: «Habría que señalar que Popper fue el primero en utilizar el término 'grado de confirmación' —hace veinte años— y que, por tanto, es desafortunado que en años recientes se haya

[6] La traducción de Carnap se encontrará, con referencias a *L. d. F.*, en su «Testability and Meaning», *Philosophy of Science* 3, 1936, p. 427. Uno o dos años más tarde sugerí a Carnap que «grado de corroboración» tal vez fuera una traducción mejor, pero como él no opinaba así, acepté su traducción. (El término «grado de corroboración» me fue sugerido por mi amigo el profesor Hugh N. Parton.)

utilizado con un sentido que no es el que le dio Popper. Pero Popper no parece darse cuenta de que el uso reciente *ha* sido en un sentido diferente»[7].

Fue esta observación de Kemeny la que me indujo a abandonar el término «grado de confirmación». Porque, evidentemente, yo no lo consideraba como pripiedad privada mía.

Pero me sorprendió un poco oír que yo «no parecía darme cuenta de que el uso reciente *ha* sido en un sentido diferente» cuando fue precisamente ese hecho obvio lo que motivó mi nota. Y los usuarios recientes del término no parecían percatarse de la posibilidad de que *el grado en el que una hipótesis ha sido contrastada puede no satisfacer el cálculo de probabilidades;* ni del hecho de que yo lo había dicho hacía mucho tiempo; ni de que mis argumentos no habían recibido respuesta.

Pero como ahora he escogido una etiqueta totalmente nueva, creo que lo mejor es ignorar la historia de la vieja que he abandonado. Lo único que importa ahora es que no vuelvan a confundirse las cosas.

En lo que concierne al «grado de confirmación» —en su sentido más reciente, que hace de él una probabilidad— no puedo dejar de sentir que el término es suplerfluo. ¿Por qué no seguir con «probabilidad»? Además ha creado mucha confusión. Porque probabilidad y «grado de confirmación» (en su uso reciente) no cumplen. Dan a todas las leyes universales, tanto si están refutadas como bien corroboradas, el mismo grado cero de probabilidad o confirmación (*cf. L. I. C.* sección 80). Y llevan a otras dificultades insuperables del mismo tipo, como descubrió Carnap posteriormente.

Por cierto, prefiero la etiqueta «grado de corroboración» a «grado de confirmación». Porque «confirmación» puede sugerir fácilmente una idea equivocada. Contiene la raíz «firm-» y sugiere o bien el proceso de hacer una hipótesis más cierta gradualmente o incluso el proceso de hacerla finalmente segura[8]. En otras palabras, el *término «confirmación»* tiene fuertes *connotaciones verificacionistas.* Por tanto, se lo entrego de muy buena gana a los verificacionistas y a los creyentes en la inducción.

Si el «grado de corroboración» se convirtiera también en una probabilidad en el trascurso de los próximos veinte años o así[9], yo sugeriría a mis discípulos (si los tengo) que utilicen,

---

[7] *Journal of Simbolic Logic* 20, 1955, p. 304 (cursiva en el original). En la réplica de Kemeny, debe hacerse la siguiente corrección, p. 304, línea 16 empezando por abajo: en lugar de «medida del apoyo dado por *y* a *x*» léase «medida del poder explicativo de *x* respecto a *y*».

[8] *Cf. Philosophy of Science* 12, 1945, nota 2 en p. 99.

[9] [Véase Jaakko Hintikka, «Induction by Enumeration and Induction by

durante otra década, más o menos, otro término sencillo, tal como «grado de una hipótesis», en lugar de «grado de corroboración» y «calificado mediante contrastaciones» en lugar de «corroborado».

30. EL PROBLEMA DEL GRADO DE CORROBORACION

Si lo tomamos como un *problema puramente práctico,* el problema de la corroboración puede reformularse como sigue: «¿Cuando aceptamos —provisionalmente— una teoría?». Nuestra respuesta, naturalmente, es: «cuando ha resistido las críticas, incluyendo las contrastaciones más severas que podemos imaginar; sobre todo cuando las ha resistido mejor que cualquier teoría rival».

Esto es todo: no es necesario llevar más lejos el problema, aunque, una vez que lo hemos resuelto, podemos plantear otro, uno que no es necesario plantear y que tiene poco que ver con la práctica de la investigación. Se trata del problema: «¿podemos evalurar el grado en el que una teoría ha resistido las contrastaciones? En particular, ¿podemos comparar dos teorías —por ejemplo, las teorías de Newton y de Einstein sobre la gravitación— y decir, con precisión, por qué la de Einstein está mejor contrastada o mejor corroborada y es, por tanto (provisionalmente), más aceptable?»

En su forma más sencilla, este problema puede expresarse así: «¿Cómo está de bien contrastada la teoría *t*?». Y «¿Podemos adjudicarle un número —una medida o grado— que resuma la severidad y el número de contrastaciones que ha sufrido la teoría y la forma en que las ha resistido, dándole algo parecido a una nota de examen, a la que podamos llamar 'grado de corroboración'?».

Expuse las propiedades características del grado de corroboración en el capítulo X de *L. I. C.* Los principales puntos de mi exposición eran éstos:

1)» El grado de corroboración está estrechamente relacionado con la *contrastabilidad* de una teoría. El hecho mismo es bastante obvio: una teoría más contrastable es una teoría que puede contrastarse mejor y lo que estamos buscando es una calificación o grado que exprese lo severamente que se ha contrastado la teoría y cómo ha resistido sus contrastaciones.

2) Puesto que la contrastabilidad puede, a su vez, medirse por medio del *contenido* de la teoría, y puesto que el contenido,

---

Elimination», en I. Lakatos, ed., *The Problem of Inductive Logic,* 1968, pp. 191-216. *Ed.*]

a su vez, puede medirse por medio de la *improbabilidad* lógica de la teoría, contenido e improbabilidad tienen una relación tan estrecha con el grado de corroboración como la propia contrastabilidad. (Puedo mencionar, de paso, que la idea del contenido empírico de una teoría, como medida de la clase de sus falsadores, fue quizá la idea *lógica* más importante de *L. I. C.* Juega un papel decisivo en la teoría de los grados de contrastabilidad, de simplicidad, de probabilidad e improbabilidad lógicas, y de corroboración.)

3) La corroboración no puede ser de ningún modo una probabilidad, puesto que está más estrechamente relacionada con la improbabilidad que con la probabilidad de una teoría: una teoría fuerte (como la teoría de las ondas electromagnéticas de luz de Maxwell) puede contrastarse más amplia y severamente que una teoría más débil entrañada por ella (tal como la teoría ondulatoria de la luz de Fresnel). Toda contrastación de la última teoría constituye una contrastación de la primera, pero lo contrario no ocurre. La situación puede ser similar incluso en el caso de teorías que estén lógicamente relacionadas de un modo algo diferente, tal como ocurre con la de Einstein y la de Newton: una contrastación que sostenga la teoría de Newton, sostiene también la de Eisntein, aunque, de hecho, algunas contrastaciones son contrastaciones cruciales entre ambas teorías; pero, además, hay contrastaciones de la teoría de Einstein que, sencillamente, no son contrastaciones de la de Newton. (Por ejemplo, la desviación del rojo en un campo gravitatorio fuerte.) En todos estos casos, es la teoría lógicamente más improbable la que no solamente es de mayor contenido o contrastabilidad, sino que pronto será, de hecho, la mejor contrastada, siempre que no resulte refutada. (Es también la que tiene mayor *poder explicativo*.)

4) El grado de corroboración de la teoría más débil (tal como la de Fresnel o la de Newton) puede no sólo ser excedido por el de una teoría fuerte (la de Maxwell o la de Einstein), incluso sin la falsación de la teoría más débil, sino que, de hecho, el grado de corroboración de la teoría más débil puede disminuir al surgir una teoría más fuerte[10].

5) Todo el mundo ha dado por sentado, acríticamente, que una hipótesis de alta probabilidad es algo bueno, algo a lo que debemos aspirar. Pero la mayor probabilidad será la de una hipótesis que no diga nada (como una tautología) o casi

[10] Esto se enuncia brevemente en la última oración antes del penúltimo párrafo de la sección 82 de *L. I. C.* Para justificar esta observación, el grado de corroboración debe definirse de tal modo que sea relativo no sólo a la evidencia en forma de contrastaciones, sino también a otras teorías. Esto se consigue por medio de lo que yo llamo su «forma relativizada»: véase *infra*, sección 32.

nada (como algunos enunciados puramente existenciales), o una que se aleje lo menos posible de los hechos que debe explicar (esto es, una hipótesis que sea *ad hoc*). No sólo no se ha examinado nunca críticamente el pretendido objetivo de obtener probabilidades altas, sino que puede mostrarse que el principio intuitivo de que las probabilidades altas son algo bueno choca con otro principio intuitivo: el principio de que las hipótesis *ad hoc* son algo malo. Y es éste último principio el que se adopta en las discusiones realmente críticas de las teorías científicas, así como en la práctica científica, y no el primero.

Hay algunos puntos más, pero los cinco que se han resumido aquí serán suficientes, espero, para mostrar lo siguiente. Mi problema real en relación con el grado de corroboración *no* era dar «una definición adecuada» de una idea intuitiva del grado de corroboración. El problema real que tenía en mente era más bien éste: hay una idea intuitiva que puede parecer a primera vista una probabilidad y que muchos lógicos han considerado como tal sin examinarla; al mismo tiempo, si miramos más de cerca nuestra evaluación real de las teorías, esta idea presenta propiedades que son incompatibles con las reglas del cálculo de probabilidades. Llegamos así a las dos cuestiones siguientes:

La primera nos plantea las preguntas ¿quién tiene razón? ¿Los filósofos que dicen que contrastando una hipótesis establecemos su probabilidad, en el sentido del cálculo de probabilidades? ¿O yo, que digo que sea lo que fuere lo que establezcamos mediante la contrastación rigurosa de una hipótesis no puede ser, en general, una probabilidad en el sentido de ese cálculo?

La segunda cuestión puede expresarse con las siguientes preguntas: ¿Es coherente mi idea del grado de corroboración? ¿Existe una función de medida que tenga las propiedades que yo atribuyo al grado de corroboración (y que, en consecuencia, no satisfaga el cálculo de probabilidades)?

Nuestro método para resolver estas cuestiones será dar una definición del grado de corroboración *en términos de contenido*, del contenido de la teoría y de los enunciados de contrastación. Como el contenido es a su vez definible en términos de improbabilidad y, por tanto, de probabilidad, nuestra definición del grado de corroboración se expresará en los términos más familiares de la probabilidad. De esta definición podemos obtener inmediatamente la solución a nuestras dos cuestiones. Dado que el *definiens* será una función de probabilidades, podemos definir al instante la cuestión de si el «grado de corroboración» es, él mismo, una probabilidad. La respuesta es negativa. Y como sabemos que el cálculo de probabilidades es coherente,

podemos contestar a la segunda pregunta afirmativamente.

El hecho mismo de que yo dé una definición del grado de corroboración hace necesario resaltar aquí que mi problema real —el que quería resolver— *no* era definir el grado de corroboración. Más bien, la tarea de encontrar una definición adecuada surgió, no de un deseo de definir, sino de algunos problemas reales. En último término, surgió de los problemas de la inducción la demarcación. Los problemas auténticos no pueden resolverse con definiciones, aunque las definiciones pueden contribuir, a veces, a clarificar ciertas cuestiones.

31. LA CORROBORACION

Lo molesto de las personas —de las personas sin sentido crítico— que sostienen una teoría, es que tienen una tendencia a considerar que todo apoya o «verifica» su teoría y a considerar que nada la refuta. Muchos empiristas han visto este peligro, por ejemplo Bacon, y han tratado de contrarrestarlo aconsejando al científico que se abstuviera de teorizar y que echase de su mente todas las teorías «preconcebidas», hasta que, como consecuencia de la observación pura, libre de prejuicios, una teoría se impusiera en su mente. Como hemos visto, este consejo es impracticable y sólo puede llevar al auto-engaño y al hábito de tener las propias teorías inconscientemente (y, por tanto, acríticamente). El consejo adecuado para el científico es decirle que siempre tendrá, consciente o incoscientemente, una multitud de teorías y que hará bien en adoptar frente a ellas una actitud crítica, aunque no pueda, por regla general, ser crítico activo de más de una teoría a la vez.

Equivale a adoptar la actitud acrítica el considerar un suceso, o una observación (digamos que e) como apoyo o confirmación de una teoría o hipótesis (h, por ejemplo) siempre que e «concuerde» con h o sea un caso de h.

De este modo se cree, acríticamente, que la ocurrencia de un cisne blanco «apoya» o «confirma» la hipótesis h_1 de que todos los cisnes son blancos. Pero ¿qué pasa con la hipótesis h_2, según la cual 90 % de todos los cisnes son blancos? Evidentemente, h_2 es incompatible con la hipótesis h_1, según la cual todos son blancos; sin embargo, si ésta última hipótesis se considera «apoyada» por la ocurrencia de un cisne blanco, también debería considerarse «apoyada» la primera[11]. Tal vez pueda

[11] C. G. Hempel, que sostiene una teoría ejemplificacionista de la confirmación, también sostiene, por extraño que parezca, la teoría de que ninguna evidencia e puede confirmar dos hipótesis h_1 y h_2 si h_1 y h_2 son incompatibles. Véase la nota 13 al Apéndice *VII de *L. I. C.*

replicarse que para el apoyo de la teoría estadística h_2, hace falta, como mínimo, una muestra de diez cisnes; nueve blancos y uno no blanco. Pero si esto fuera así, ¿no deberíamos decir también que en apoyo de h_1 se necesita al menos el mismo número de cisnes, todos ellos blancos, claro? (Pero si así fuese, entonces un caso no podría sustentar nunca una teoría.)

De todos modos, pospondremos la cuestión de las teorías estadísticas, como h_2, para más tarde; por el momento consideraremos solamente las teorías universales.

Respecto a éstas, tenemos que considerar dos actitudes principales:

a) La actitud acrítica o verificacionista: uno busca una «verificación» o «confirmación» o «ejemplificación» y, por regla general, la encuentra. Todo «ejemplo» observado de la teoría se considera una «confirmación» de ella.

b) La actitud crítica, o actitud falsacionista: uno busca falsaciones o contra-ejemplos[12]. Sólo si no tiene éxito la búsqueda más concienzuda de contra-ejemplos, podemos hablar de una corroboración de la teoría.

Puede uno preguntar si es realmente tan acrítico considerar todos los ejemplos de una teoría como confirmaciones de ella. Pero puede mostrarse que, por razones lógicas, hacerlo equivale a creer que *todo es una confirmación* con la excepción de un contra-ejemplo. Así, no sólo los cisnes blancos, sino los cuervos negros y los zapatos rojos «confirman» la teoría de que todos los cisnes son blancos. Esto puede parecer raro; pero si es cierto, entonces debe admitirse que los que consideran (casi) todo como confirmación de sus teorías están actuando muy apropiadamente desde el punto de vista verificacionista y viceversa: que el verificacionista es realmente un hombre que tiende a tomarlo todo como confirmación de sus concepciones.

Es fácil mostrar que todo es, efectivamente, un ejemplo de «todos los cisnes son blancos», a menos que sea un contra-ejemplo, es decir, un cisne no blanco. Porque «todos los cisnes son blancos» puede escribirse: «cualquier cosa que sea un cisne, es blanca» o «todo es blanco a menos que no sea un cisne». Siempre que una cosa satisfaga esto, es un «*ejemplo*». Es bien sabido (en lógica extensiva) que cualquier cosa que no sea un cisne satisface esas fórmulas y también cualquier cosa que sea blanca; de modo que *sólo* una cosa que sea al mismo tiempo un cisne y no blanca (y, por tanto, un contra-ejemplo) no las satis-

[12] No carece de interés histórico el que, hasta 1600 aproximadamente, «instance» [ejemplo, en el texto] significaba «contra-ejemplo» [*counter-example, counter-instance*] (de *instantia*, que es un obstáculo: lo que se interpone en el camino, en este caso, de la aceptación).

face. Esto muestra por qué la verificación o ejemplificación es demasiado fácil para ser significativa y por qué sólo son interesantes los casos de los que se puede esperar que sean contraejemplos, a menos, claro, que la teoría sea verdadera. Más exactamente, los casos interesantes serán los casos *cruciales*: casos para los cuales la teoría a contrastar predice resultados que difieren de los resultados predichos por otras teorías significativas, sobre todo por las teorías que han sido aceptadas hasta el momento.

La diferencia fundamental entre la concepción del «apoyo» del verificacionismo y la mía es, pues, ésta: mientras que la concepción del verificacionismo lleva a la afirmación de que *todo «ejemplo»* de *h* apoya a *h*, yo mantengo que sólo las *contrastaciones auténticas* pueden apoyar a *h*.

Así, un cisne blanco observado, para el verificacionista, apoyará la teoría de que todos los cisnes son blancos; y si es coherente (como Hempel), entonces dirá que un corcorán negro observado también apoya la teoría de que todos los cisnes son blancos. En mi concepción, por otro lado, ninguna de esas observaciones apoya necesariamente esta teoría, aunque cualquiera de ellas *puede* apoyarla en circunstancias especiales. Así pues, si, por ejemplo, tenemos *buenas* razones para pensar (a la luz de teorías aceptadas previamente) que la cosa que hay en este estanque es un cisne negro, entonces el descubrimiento de que es un cisne blanco o un corcorán negro *podría* apoyar la teoría de que todos los cisnes son blancos.

En la concepción verificacionista de la cuestión, existe una relación sencilla de lógica formal —la ejemplificación— cuya presencia o ausencia decide si *e* apoya o no a *h*. Si *e* es un ejemplo de *h*, entonces *e* apoya a *h*. En mi propia concepción, la situación es menos sencilla: sólo si *e* es el resultado de auténticos y sinceros intentos de refutar *h*, puede considerarse que *e* apoya *h*.

Me parece poco probable que esta relación entre *e* y *h* sea susceptible de ser analizada lógicamente por completo: la sinceridad no es el tipo de cosa que se presta al análisis lógico. Pero no hay razón para descorazonarse. En primer lugar, podemos arreglárnoslas bastante bien sin un análisis lógico del apoyo y de la corroboración. Y en segundo lugar, podemos llegar bastante cerca de su análisis, más cerca de lo que podría esperarse.

Trataré ahora de desarrollar más plenamente la idea «*e* apoya a *h*», en preparación del análisis del grado en el cual *e* corrobora a *h*. Con este fin utilizaremos, además de *e* (la evidencia empírica) y *h* (la hipótesis), una nueva variable: nuestra

información básica, *b*, y podremos decir entonces que *e* apoya a *h* a la luz de *b*[13].

Por información básica *b* entendemos cualquier conocimiento (pertinente a la situación) que aceptamos —quizá sólo provisionalmente— mientras contrastamos *h*. Así, *b* puede, por ejemplo, incluir condiciones iniciales. Es importante advertir que *b* debe ser coherente con *h*; por tanto, si antes de considerar y de contrastar *h*, aceptásemos una teoría *h'* que, junto con el resto de nuestra información básica fuese incongruente con *h*, entonces tendríamos que excluir *h'* de la información básica *b*. (Puede que no siempre sea fácil decidir qué es lo que hay que excluir de *b*, pero éste es un problema que no quiero discutir aquí.)

Supongamos que hay unas consecuencias experimentalmente contrastables de *hb*, es decir, de la conjunción de *h* y *b*. Entre esas consecuencias puede haber algunos sucesos que debemos suponer que no ocurrirán si *h* es falsa. Sea *e* una de esas consecuencias, es decir, sea *e* la descripción de un suceso observable del que puede predecirse que ocurrirá (en presencia de *b*) si *h* es verdadera y que debemos esperar que no ocurra si *h* es falsa. En ese caso, tenderemos a decir que si *e* ocurre de hecho, entonces esto apoya a *h*.

Sin embargo, ésta no es una formulación muy satisfactoria, aunque puede servir de punto de partida. El problema está en la advertencia (hecha dos veces) de que *e* es un suceso «que debemos esperar que no ocurra si *h* es falsa».

En vista del hecho de que exigimos que «*e* se siga de *h* (en presencia de *b*)» uno podría sentirse inclinado a interpretar la advertencia en cuestión como la exigencia de que «no-*e* se siga de no-*h* (en presencia de *b*)». Pero esto equivaldría a otra forma de verificacionismo: haría a *h* y a *e* equivalentes (en presencia de *b*) y nos permitiría verificar *h* por medio de la observación, es decir, al observar que *e* es verdadero.

Pero, aparte de cualquier hostilidad hacia el verificacionismo, es totalmente implausible exigir que no-*e* se siga de no-*h* (y *b*). Porque supongamos que *e* es un suceso que apoya a *h*, algo predicho por *h* y algo que nadie hubiera tomado en consideración nunca sin *h*. Por ejemplo, sea *e*, la primera observación de un nuevo planeta (Neptuno) por J. G. Galle, en una posición predicha por Adams y Leverrier, y sea *h* la teoría de Newton sobre la cual se basaba esta predicción. Entonces *e* apoya efectivamente a *h*, y muy fuertemente. Pero, a pesar de este hecho, *e* también se sigue de teorías que, como la de

[13] [La idea de «información básica» que introduce Popper aquí, fue publicada por él por primera vez en 1954 y 1957. Véase *L. I. C.*, pp. 369, 372-3. *Ed.*]

Einstein, entrañan no-*h* (en presencia de *b*). Sería, por tanto, un error de gran importancia exigir que no-*e* se siguiera de no-*h* (y *b*). El error equivale a creer que *e* (que, hemos dicho, debe ser crucial entre *h* y *alguna* otra teoría significativa) tiene que ser crucial entre *h* y *todas* las demás teorías que contradicen a *h*.

Por esta razón, nuestra segunda exigencia «no-*e* se sigue de no-*h* (y *b*)» es demasiado fuerte. Pero no podemos pasarnos por completo de algo de este tipo. Porque nuestra primera exigencia, que *e* se siga de *h* (y *b*), es totalmente insuficiente, considerando que *e* se sigue de *h*, en presencia de *b*, si se sigue de *b* sola.

Pero tampoco es suficiente exigir que *e* no se siga de *b* sola. Porque si *e* fuese *probable* en presencia de *b* sola («probable» en el sentido del cálculo de probabilidades), entonces su ocurrencia difícilmente podría considerarse como un apoyo significativo de *h*.

Así, James Challis, a quien Adams había dado el resultado de sus cálculos, observó realmente a Neptuno cerca de la órbita calculada, antes que Galle. Pero la estrella que vio no parecía moverse y no pensó que su observación tuviese una importancia suficiente para compararla con otras observaciones posteriores de la misma región que hubiesen revelado su movimiento. La presencia de *alguna* estrella desconocida de la octava magnitud cerca de la situación calculada era, en sí misma, bastante probable dada la información básica que él poseía y, por tanto, no le pareció significativa. Sólo la presencia de una estrella *en movimiento,* un planeta, hubiera sido significativa, porque era inesperada, aunque no según los cálculos de Adams[14].

De modo que supondremos que lo inesperado de un suceso puede identificarse con una *baja probabilidad,* en el sentido del cálculo de probabilidades, dada la información básica.

Es, pues, necesario plantear una segunda exigencia; por medio de ella podemos excluir, por ejemplo, aquellos casos en los que *e* sería probable, dada, solamente, la información básica *b*, «probable» en el sentido del cálculo de probabilidades.

Necesitamos, por tanto, el cálculo de probabilidades; y escribiremos

$$p\,(a,b) = r,$$

que debe leerse: «La probabilidad de *a* dada *b* es igual a *r*».

[14] Para los hechos, véase B. A. Gould, *Report to the Smithsonian Institute on the History of the Discovery of Neptune*, 1850.

Del mismo modo escribiremos «*ab*» para la conjunción de «*a* y *b*» y «$\bar{a}$» para «no-*a*». En consecuencia.

$$p\,(ab,c) = r$$

debe leerse: «La probabilidad de *a* y *b*, dada *c*, es igual a *r*»; y, por ejemplo,

$$p\,(\bar{a},bc) = r$$

debe leerse: «La probabilidad de no-*a*, dadas *b* y *c*, es igual a *r*».

Ahora podemos expresar nuestra exigencia de que la evidencia empírica *e*, si ha de apoyar a *h*, no sea probable (o esperada) sobre la sola base de la información básica *b* por

$$p\,(e,b) \ll 1/2$$

Esto nos lleva a percatarnos al instante de que *cuanto menor sea p (e,b), mas fuerte* será el apoyo que *e* preste a *h*, siempre que *e* se siga de *h* y *b* o de *h* en presencia de *b*.

Este resultado es muy interesante de dos formas. En primer lugar, intuitivamente es muy satisfactorio que sólo una evidencia improbable *e* —improbable dada nuestra información básica *b*— sea aceptada como significativa, como lo muestra nuestro ejemplo sobre Neptuno. La predicción de un adivino a una señorita «pronto conocerá usted a un interesante joven» es demasiado probable en sí misma para ser aceptada, si se realiza, como apoyo de las teorías sobre las que puede basarse el arte de la adivinación. En segundo lugar, exigimos intuitivamente que sólo cuenten las contrastaciones *severas* y que cuanto más severas sean, más cuenten. Pero esto es lo mismo que exigir que *e* sea improbable dada nuestra información básica. Esto expresa claramente en lenguaje probabilístico la exigencia de que *e* sea el resultado de una contrastación *severa*. Y esto, a su vez, implica la exigencia de que *h* sea sumamente contrastable —es decir, contrastable por medio de contrastaciones *severas*— y, por tanto, sumamente improbable o de gran contenido empírico.

Todas estas consideraciones sugieren que la siguiente definición de «*e* apoya a *h*» sería razonable y adecuada:

e apoya a *h* en presencia de la información básica *b* si, y sólo si,

a) *e* se sigue de *hb*

b) $p\,(e,b) \ll 1/2$

Es evidente que podemos escribir *a)*, como *b)*, en términos probabilísticos; es decir, en vez de «*e* se sigue de *hb*» podemos

utilizar la fórmula ligeramente más débil[15]: «la probabilidad de *e* dadas *hb* (es decir, la conjunción de *h* y *b*) es igual a uno». Lo escribiremos así:

a) $$p\,(e,\ hb) = 1$$

En vista de *a)* y *b)*, podemos modificar (y debilitar) aún más nuestra definición escribiendo:

e apoya a *h* en presencia de *b* si, y sólo si,

c) $$p\,(e,\ hb) - p\,(e,b) > 1/2$$

Puede mencionarse entre paréntesis que, escrito de esta forma, «apoya» resulta ser una versión más fuerte de «depende positivamente de»; porque definimos, generalmente, en teoría de la probabilidad:

a depende positivamente de *b* en presencia de *c* si, y sólo si,

$$p\,(a,bc) - p\,(a,c) > 0$$

Y, de hecho, podríamos usar esto como una definición de «apoya» o, más exactamente, podríamos sustituir el *definiens c)* por el *definiens* más débil

d) $$p\,(e,\ hb) - p\,(e,\ b) > 0$$

De este modo llegamos a la forma (relativizada) de la definición de «apoyo» dada en *L. I. C.*, Apéndice *IX en la primera de las tres notas (p. 368).

Me parece a mí que el *definiens d)* es bastante adecuado siempre que le añadamos el siguiente aditamento que, por así decirlo, sustituye a *c):*

El apoyo dado por *e* a *h* se hace *significativo* sólo cuando

e) $$p\,(e,\ hb) - p\,(e,\ be) \gg 1/2$$

Con este aditamento hemos entrado en el problema del *grado* de apoyo o *grado* de corroboración. Porque al aceptar el *definiens d)* y añadir este aditamento decimos que no es suficiente definir «apoya» sino que tenemos, más allá de esto, que determinar un *grado significativo* de apoyo; o, en otras palabras, que tenemos que introducir *una medida del grado de apoyo o de corroboración* para distinguir grados significativos.

---

[15] «*e* se sigue de *hb*» puede representarse en lenguaje probabilístico, por ejemplo por la fórmula

$$p(e,hb\bar{e}) = 1$$

De esta fórmula se sigue inmediatamente la fórmula *a*) más débil. Véase *L. I. C.*, Apéndice *V, la fórmula (*) en la página 332 y la discusión que la sigue.

La primera sugerencia, obvia, parece ser ésta: al aceptar el aditamento, tomamos la diferencia usada en *c)* y *d)*, es decir,

$$f) \qquad p\,(e,\ hb) - p\,(e,b)$$

como medida del grado de apoyo dado por *e* a *h* en presencia de *b* o del grado de corroboración *(C)* de *h* por *e* en presencia de *b*, es decir, la sugerencia es que definamos

$$C\,(h,\ e,\ b),$$

esto es, el grado de corroboración de *h* por *e* en presencia de la información básica *b*, por medio de la diferencia expresada en la definición *f)*. La definición sugerida tiene algunas imperfecciones que pueden subsanarse «normalizando» *e)*, esto es, dividendo *e)* por el «factor de normalización»

$$g) \qquad p\,(e,\ hb) - p\,(eh,\ b) + p\,(e,b)$$

Llegamos así a la definición

$$\text{D} \qquad C\,(h,\ e,\ b) = \frac{p\,(e,hb) - p\,(e,b)}{p\,(e,hb) - p\,(eh,b) + p\,(e,b)}$$

Mientras en esta definición, D, el numerador *e)* de la fracción tiene una significación intuitiva clara y sencilla, el denominador *f)* no tiene tal significación: se escoge simplemente porque lleva a resultados satisfactorios —elimina las imperfecciones mencionadas— y porque parece ser el factor de normalización más sencillo entre los que llevan a esos resultados.

Antes de señalar una de las tachas que elimina la adopción de D mencionaré algunos de los resultados generales de adoptar *f)* o D.

Tanto si D se adopta como definición del grado de corroboración como si se escoge *e)* como *definiens*, obtenemos los siguientes resultados, sumamente intuitivos:

i) Si *e* apoya a *h* (dada la información básica *b*), entonces *C (h,e,b)* es positiva. Si *e* mina a *h* (de modo que no-*e* apoya a *h*), entonces *C (h,e,b)* es negativa. Si *e* no hace ninguna de las dos cosas, de forma que es *independiente* de *h* en presencia de *b*, entonces *C (h,e,b)* es igual a cero.

Por ejemplo, si *e* es una tautología, o una consecuencia lógica de *b*, entonces *e* no corroborará ni minará a ninguna teoría *h* y *C (h,e,b)* será cero. Del mismo modo, si *h* es una tautología o una consecuencia lógica de *b*, no será ni corroborada ni minada por ninguna evidencia *e* y, una vez más, $C\,(h,e,b) = 0$.

ii) El valor máximo que puede alcanzar *C (h,e,b)* es 1. Este valor se alcanza si, y sólo si,

aa) $p\ (e,hb) = 1$

bb) $p\ (e,b) = 0$

Este resultado es sumamente satisfactorio: sólo una *e* que sea extremadamente improbable dada la información básica *b* puede dar a *h* el máximo apoyo o un grado máximo de corroboración. Desde luego, para hacer esto, *e* debe también, en presencia de *b* seguirse (o «casi seguirse») de *h*, esto es, *p (e,hb)* debe ser igual a 1. Finalmente, para que *h* sea corroborable en grado máximo, su contenido o grado de contrastabilidad debe ser máximo; es decir, la siguiente fórmula es una condición más que debe ser satisfecha para que la corroboración alcance su valor máximo 1:

cc) $p\ (h,b) = 0$

En efecto, *cc)* es una consecuencia inmediata de *aa)* y *bb)*[16].

iii) Más en general, incluso si $p\ (h,b) \neq 0$, el valor máximo que puede alcanzar *C (h,e,b)* es igual a $1 - p\ (h,b)$ y, por tanto, igual al contenido de *h* en relación con *b*, o a su grado de contrastabilidad. Esto hace al grado de contrastabilidad igual al máximo grado de corroboración de *h*, o a su «grado de corroborabilidad».

Hasta aquí, todo bien. Pero ahora veamos una imperfección: ¿Qué pasa con una evidencia empírica *e* que false *h* en presencia de *b*? Tal *e* hará *p (e,hb)* igual a cero. Pero si esta *e* comunica el resultado de una contrastación muy severa —por ejemplo, una medición muy precisa— entonces puede muy bien ocurrir que sea muy improbable en relación con *b*; es decir, que no sólo tenemos $p\ (e,hb) = 0$ o sino, aproximadamente, $p\ (e,b) \approx 0$. En este caso, el grado de apoyo o corroboración se hará cero (o muy poco menos) si adoptamos *f)* como *definiens*. Pero si adoptamos D, entonces el grado de apoyo o de corroboración será siempre igual a — 1 cuando *e* false a *h* (en presencia de *b*); es decir, obtenemos, a partir de D:

Si *e* es incompatible con *hb*, entonces

$$C\ (h,e,b) = -1$$

Esto es claramente satisfactorio y elimina un defecto inherente en la elección del *definiens f)*. Hace, para toda *h* (siempre

[16] *cc)* se sigue de *aa)* y *bb)* en presencia de las fórmulas B1 y B2 de *L. I. C.*, Apéndices *IV y *V:

$$p(e,b) \geq p(eh,b) = p(e,hb)\ p(h,b).$$

que sea coherente con *b*) los grados máximo y mínimo de corroboración iguales a — 1 y al contenido o grado de contrastabilidad de *h* (cuyo máximo es + 1).

(Otros aspectos satisfactorios se mencionan con cierta extensión en las tres notas reimpresas en el Apéndice *IX de *L. I. C.*; véanse sobre todo los nueve «*desiderata*» enunciados allí en la sección 9 de la primera nota.)

Hay definiciones distintas de D y algunas de ellas pueden ser preferibles a D desde algunos puntos de vista. De hecho, mi primera definición[17] era ligeramente diferente a D. Prefiero D porque me parece la más sencilla y la más lúcida de las diversas fórmulas que satisfacen mis *desiderata*. Pero ciertas fórmulas logarítmicas pueden ser igual de buenas o mejores para ciertos propósitos.

Lo importante de todas estas fórmulas que son satisfactorias es esto: todas ellas parecen ser «topológicamente equivalentes»; es decir, si adoptando D tenemos, para algunas *h*, *e*, *b* y *h'*, *e'*, *b'*, que

$$C\,(h,e,b) \geq C\,(h',e',b')$$

entonces la misma relación es válida si adoptamos cualquiera de las demás definiciones satisfactorias. Pero esto significa que filosóficamente, no hay nada que obligue a escoger entre ellas, excepto desde el punto de vista de la simplicidad y la perspicuidad.

Pero, para todas las definiciones «*satisfactorias*» —es decir, para todas las que son topológicamente equivalentes a D— se cumple el siguiente teorema:

El grado de corroboración no es una probabilidad, es decir, no satisface las reglas del cálculo de probabilidades[18].

Esto ya se ha aceptado ampliamente, aunque hay algunas personas que siguen insistiendo en que la «aceptabilidad», al contrario que la «corroboración», es una probabilidad. Naturalmente, podemos definir así la «aceptabilidad» (o «confirmación»). Pero si por grado de aceptabilidad de una teoría *h* entendemos el grado en que *h* es satisfactoria desde el punto de vista de la *ciencia empírica* —es decir, desde el punto de vista de *los objetivos de la ciencia empírica*— entonces la aceptabili-

---

[17] Véanse las fórmulas 9.2 y 10.1 en *L. I. C.*, Apéndice *IX, primera nota (pp. 368-369). La actual definición D se encuentra allí como (10.1*) en la nota *2 (p. 372).

[18] Lo mismo es válido para *l(h,e)* la «verosimilitud de *h*» en el sentido de Fisher, definida por

$$l(h,e) = p(e,h);$$

porque aunque es una probabilidad, no es una de *h*.

dad se habrá convertido en topológicamente equivalente a la corroboración. Las tautologías tienen el grado máximo de probabilidad, pero son inaceptables para la ciencia empírica: establecer que una supuesta teoría empírica es tautológica (o casi) siempre equivale a una crítica decisiva de ella.

Así pues, no niego que pueda decirse que la interpretación lógica de la probabilidad, o la probabilidad de los enunciados, da el grado de probabilidad o de verosimilitud de que un enunciado sea verdadero; y en la medida en que lo hace, da el grado de «aceptabilidad», *siempre que* al llamar «aceptable» a un enunciado, queramos decir que creemos que es verdad. Pero esto no es todo lo que queremos decir cuando aceptamos un enunciado en ciencia. Esto puede verse fácilmente en el hecho de que las teorías científicas interesantes tienen siempre una probabilidad despreciable (si no es cero), incluidas aquellas que son generalmente aceptadas actualmente.

Concluyo esta sección dando un resumen de mis concepciones respecto a la corroboración, en forma de siete puntos, el primero de los cuales contiene la idea fundamental.

1) El grado de corroboración de una teoría es una evaluación de los resultados de las contrastaciones empíricas a las que se ha sometido.

2) Hay dos actitudes, dos modos de mirar a las relaciones entre una teoría y la experiencia: uno puede buscar confirmación o refutación. (Estas dos actitudes son, obviamente, variantes de la actitud apologética [o dogmática] y de la actitud crítica.) *Las contrastaciones científicas son siempre intentos de refutación.*

3) La diferencia entre los intentos de confirmación y los intentos de refutación o contrastaciones es, en gran medida, aunque no totalmente, susceptible de análisis lógico.

4) Se dirá que una teoría está mejor corroborada cuanto más severas sean las contrastaciones que ha superado (y cuanto mejor las haya superado).

5) Se dirá que una contrastación es tanto más severa cuanto mayor sea la probabilidad de fracasar en ella (la probabilidad absoluta o previa y también la probabilidad a la luz de lo que llamo nuestra «información básica», es decir, el conocimiento que, por común acuerdo, no se pone en duda mientras se contrasta la teoría bajo investigación).

6) Así, toda contrastación auténtica puede describirse, intuitivamente, como un intento de «pillar» a la teoría: no sólo es un examen severo, sino que, como examen, es injusto: se emprende con el objetivo de suspender al examinado, más que con el objetivo de darle la oportunidad de mostrar lo que sabe.

Esta última actitud sería la de un hombre que quisiera confirmar o «verificar» su teoría.

7) Suponiendo que siempre nos guía, en nuestras contrastaciones, una actitud auténticamente crítica y que *nos esforzamos por contrastar la teoría* (una suposición que no puede formalizarse)[19], podemos decir que el grado de corroboración de una teoría aumentará con la improbabilidad (a la luz de la información básica) de las predicciones que constituyen los enunciados contrastadores, siempre que las predicciones derivadas con ayuda de la teoría tengan éxito[20].

32. ALGUNOS COMENTARIOS MAS SOBRE LA DEFINICION DEL GRADO DE CORROBORACION

Formularemos en esta sección algunas nuevas condiciones o exigencias o requisitos que debe satisfacer el grado de corroboración *C (h,e,b)*. (Véase también la lista de *desiderata* en *L. I. C.*, pp. 372 ss.) Estas condiciones y, sobre todo, el problema de si son coherentes, me llevó originariamente a la definición *C (a,b)* y más tarde al *C (a,b,c)* o, como la escribí en la sección 31, *supra*, a *C (h,e,b)*.

1) El principal requisito es que *nuestro concepto debe darnos una estimación del grado en el cual una hipótesis o teoría ha sido contrastada.* Casi todas las restantes exigencias se derivan de esta exigencia central (aparte de las exigencias de carácter convencional).

2) Cuanto mejor, más severamente, pueda contrastarse una teoría, mejor podrá corroborarse. Por tanto exigimos que *contrastabilidad y corroborabilidad crezcan y decrezcan juntas*, por ejemplo, que sean *proporcionales*. Esto haría a la corroborabilidad inversamente proporcional a la probabilidad lógica (absoluta); véase *L. I. C.*, secciones 34, 35, 83[21]. (Por «corrobo-

[19] «Aun cuando es posible que la lógica establezca criterios para decidir si un enunciado es contrastable, en ningún caso se ocupa sobre si nadie se esfuerza o no por contrastarlo.» (Véase *L. I. C.*, sección 11, p. 52.)

[20] [Aquí parecen faltar algunas líneas del texto y no pudieron ser reconstruidas. *Ed.*]

[21] Véase el Apéndice *IX de *L. I. C.* donde se reimprimen mis tres artículos sobre «El grado de corroboración» (*Brit. Journ. Phil. Science* 5, 1954, pp. 143 ss. Véanse también las correcciones en 5, pp. 334, y 359; 6, 1955, pp. 157 ss.; 7, 1957, pp. 244 s., 249 ss. y 350 ss.) Nótese que en estos artículos usé el término «confirmación» en lugar de «corroboración» y que usé el símbolo «*C(x)*» para el contenido de *x*, mientras que aquí uso el símbolo «*Ct(x)*». En el primero de esos artículos se enuncian varios *desiderata* que no menciono aquí. Véanse también los Apéndices *II y IV. [Véanse también los Addenda a *El desarrollo del conocimiento científico (C. & R.) Ed.*]

rabilidad» entiendo el grado más alto de corroboración que pueda obtener teóricamente la hipótesis en cuestión.)

3) La convención más sencilla consistirá en suponer que el factor de proporcionalidad es igual a 1; o, en otras palabras, que *la corroborabilidad es igual a la contrastabilidad y al contenido empírico.*

4) Una tautología —un enunciado analítico— no puede contrastarse empíricamente. Esto puede expresarse mediante la exigencia (que se sigue de la anterior) de que su grado de corroboración sea cero, cualquiera que sea la evidencia y, en consecuencia, también lo será su grado de corroborabilidad.

5) Una hipótesis auto-contradictoria tiene, naturalmente, un contenido máximo; por tanto, tienen una contrastabilidad máxima, pero sólo en el sentido de que toda contrastación puede ser considerada como una refutación. Podemos abordarla o bien excluyendo las hipótesis auto-contradictorias o mediante la exigencia de que una hipótesis auto-contradictoria tenga siempre, cualquiera que sea la evidencia, ese grado mínimo de corroboración que es indicativo de una *refutación.*

6) En vista de 3) y de 4), éste no puede ser cero, y por ello tenemos que tomar números negativos para indicar los resultados desfavorables. Una convención simple consiste en usar el número — 1 para caracterizar el grado de corroboración de una hipótesis con respecto a las contrastaciones que la refutan.

7) Esto es aconsejable ya que, en vista de la convención de hacer el factor de proporcionalidad igual a 1, formulada en el punto 3), el grado máximo de corroboración que puede tener incluso la hipótesis más contrastable, es + 1. Obtenemos así límites simétricos, — 1 y + 1, para el grado de corroboración, mientras que 0 significa ausencia de corroboración, como en el caso de una tautología.

8) Si introducimos ahora el símbolo «*C (a,b)*» para el grado de corroboración de *a* y *b* y «*Ct (a)*» para el contenido empírico o grado de contrastabilidad de *a*, de tal modo que

$$Ct\,(a) = 1 - p\,(a),$$

podemos expresar nuestras exigencias, tal como se han formulado hasta el momento, por

$$\text{i)} \qquad -1 \leq C\,(a,b) \leq Ct\,(a) \leq +1$$

Nótese que «*b*» *no* se usa aquí para «información básica» que es, más adelante, en la sección, indicada por «*e*».

9) Si escribimos «$\bar{a}$» con el significado de la negación de *a*, podemos decir que $\bar{a}$ refuta a *a*. sí pues, $C\,(a,\bar{a}) = -1$.

Del mismo modo, si escribimos «*ab*» para la conjunción de *a* y *b*, $C(a\bar{a},b) = -1$, al menos cuando *b* sea coherente, puesto que $a\bar{a}$ es un enunciado contradictorio. Además, un enunciado *a* —por ejemplo, una hipótesis singular— será completamente corroborada si la evidencia misma entraña a *a*. Así, $C(a,a) = C(a,ab)$ será el máximo grado de corroboración del que es capaz un *a* coherente, es decir que $C(a,a)$ será igual a lo que hemos llamado la corroborabilidad de *a* y así, a su contenido $Ct(a)$. Si incorporamos todo esto a nuestra fórmula, admitiendo sólo la evidencia coherente *b*, obtenemos:

ii) $$-1 = C(a\bar{a},b) = C(a,\bar{a}) \leq$$

$$\leq C(a,b) \leq C(a,a) \leq Ct(a) \leq +1$$

Esto determina los límites de $C(a,b)$. Lo que es convencional aquí es simplemente la suposición: *a)* según la cual $C(a,a)$, el grado máximo de corroboración de que es capaz *a* (y, por tanto, su corroborabilidad), es igual a $Ct(a)$, su contenido, en lugar de ser solamente proporcional a él; y *b)* que el grado mínimo de corroboración (que es el grado de una hipótesis refutada) es el negativo del grado máximo que pueden obtener los enunciados más contrastables, lo que significa que es igual a -1.

10) Considerando que $\overline{a\bar{a}}$ («no [$\bar{a}$ y no-*a*]») es una tautología, deberíamos tener también para toda *b*,

$$C(\overline{a\bar{a}},b) = 0$$

11) Fijados los límites y cero, exigimos ahora que el grado de corroboración de una hipótesis aumente con cada contrastación auténtica que supere y que aumente más cuanto más severa sea la contrastación.

Suponiendo que sepamos que *b* ha sido cocebida como intento real de refutar a *a* —es decir, como una auténtica contrastación de *a*— podemos medir la severidad de la contrastación *b* por medio de la *improbabilidad* (absoluta) de *b*. Por ejemplo, las predicciones de Adams y Leverrier, que llevaron al descubrimiento de Neptuno, constituyeron una espléndida corroboración de la teoría de Newton, a causa de la extrema *improbabilidad* de que un planeta nunca observado fuese a descubrirse, por puro accidente, en la reducida región del cielo donde sus cálculos lo habían situado. Otro llamativo ejemplo es la predicción de Einstein de un eclipse, una predicción relativa a la distancia angular entre estrellas fijas que aparecen en lados opuestos del disco solar. (Einstein predijo que esta distancia

sería mayor que la que se obtuviera a partir de mediciones en el cielo nocturno.) Aquí se predijo un efecto que nadie[22] había pensado antes, y que, *a la luz de las teorías anteriores,* era sumamente *improbable.* (La relatividad general de Einstein predijo una desviación de la distancia en el cielo nocturno dos veces mayor que la que podía obtenerse a partir de la teoría de Newton.)

Estos ejemplos ilustran cómo atribuimos a predicciones arriesgadas y altamente improbables que resultan tener éxito, la capacidad de aumentar mucho el grado de corroboración. Porque en estos casos, el éxito de una predicción difícilmente puede deberse a una coincidencia o al azar: la improbabilidad de la predicción da la medida de *la severidad de la contrastación.* El argumento puede utilizarse también en la dirección contraria: el éxito de predicciones confusas y cautelosas tales como las de los astrólogos o los que echan la buena ventura, suelen producir poca capacidad de corroboración, ya que la falta de precisión corresponde a una elevada probabilidad absoluta (*cf.* sección 37 de *L. I. C.*). Este tipo de éxito puede deberse, simplemente, a una coincidencia (como se explica en la sección 83 de *L. I. C.*). El argumento también es aplicable, naturalmente, al resultado de una *multiplicidad* de contrastaciones, puesto que un enunciado que describa muchas contrastaciones (especialmente si son independientes entre sí) será menos probable que un enunciado que describa sólo algunas de esas contrastaciones. (Así, una contrastación múltiple es más improbable —y, por consiguiente, *mas severa*— que las contrastaciones que la componen.) Podemos resumir todo esto en la siguiente regla:

iv) Siempre que *b* sea una contrastación auténtica de *a*, *C (a,b)* aumenta con *Ct (b),* esto es, con el contenido, o la improbabilidad lógica absoluta, de *b*; o, dicho de otro modo, con la severidad de la contrastación *b*.

12) La exigencia que acabamos de formular tiene gran importancia para nuestra teoría en varios aspectos. Una razón de su importancia es que *no sólo es válida para el grado de corroboración, sino también para la probabilidad.* La probabilidad lógica de una hipótesis *a*, relativa a la evidencia *b* también aumenta con la improbabilidad absoluta de *b*, siempre que la probabilidad lógica absoluta de *a* no sea cero y que *b* sea deri-

[22] [* Cuando escribí esto no sabía que J. von Soldner había calculado ya en 1801 la combadura de las partículas luminosas al rozar el sol. Esto lo revela un notable artículo de Stanley L. Jaki, en *Foundations of Physics* 8, 1978, pp. 927 ss.]

vable o predictible por *a*. Podemos denotar la probabilidad relativa de *a* dada por *b* por medio de

$$p\,(a,b)$$

y la probabilidad absoluta de *a* por

$$p\,(a)$$

(Aquí $p\,(a)$ puede definirse como $P\,(a,\overline{a\bar{a}})$, es decir, como la probabilidad relativa de *a* no dando nada —o dando la tautología— como evidencia.) Podemos expresar ahora la regla del cálculo de probabilidades (una forma del teorema de Bayes) que corresponde a iv) como sigue:

iv)$_p$ Siempre que sea $p\,(a) \neq 0$ y *b* sea predictible a partir de *a*, $p\,(a,b)$ aumenta con $p\,(\bar{b})$, es decir, con la probabilidad absoluta de no-*b*, que es lo mismo que la improbabilidad lógica de *b* o $Ct\,(b)$, el contenido de *b*.

Se verá que esta regla iv)$_p$ corresponde exactamente a iv); y esto puede sugerir que el grado de corroboración puede ser, después de todo, una probabilidad, es decir, puede satisfacer el cálculo de probabilidades o quizá un cálculo muy similar[23]. En otras palabras, la analogía entre iv) e iv)$_p$ muestra la pertinencia del *primero de nuestros dos problemas: si la corroboración es, o no, una probabilidad* (*cf.* sección 30).

13) Pero esta analogía plantea el *segundo problema: si los requisitos aquí explicados son coherentes.* Porque iv) puede tomarse como indicación de que la corroboración, incluso aunque no sea una probabilidad, debe ser muy similar a una probabilidad. Pero nuestra exigencia 2) e, incluso más claramente, 4), que es la misma que iii), son *incompatibles* con una interpretación de la corroboración como una probabilidad. Porque 4) e iii) dan un grado de corroboración cero a todas las tautologías —que tienen, naturalmente, el grado más alto de probabilidad, es decir, 1—.

Por tanto, aquí es donde se hace patente la importancia de nuestras dos cuestiones (*cf.* sección 30).

14) Es fácil mostrar con detalle que nuestra exigencia 2) contradice la concepción de que la corroboración es una probabilidad; pero la exigencia 2) se sigue inmediatamente de la idea fundamental de que el grado de corroboración es simplemente un informe sobre la severidad de las contrastaciones. (Véase también la introducción que precede a la sección 79 de *L. I. C.*). Por tanto 2) y sus consecuencias, que mostrarán

[23] *Cf.* nota 1 a la sección 81 de *L. I. C.* con referencias a Janina Hosaisson-Lindenbaum.

ahora, no pueden abandonarse; lo que hace más urgente aún demostrar que nuestras exigencias son coherentes, a pesar de la analogía entre iv) e $iv)_p$.

Supongamos que el enunciado *b* se sigue del enunciado *a* pero no viceversa y que $0 \neq p\,(a) < p\,(b) \neq 1$. En este caso, habrá un enunciado *c* tal que *a* sea equivalente a la conjunción *bc* y $0 \neq p\,(c) \neq 1$.

Puesto que el contenido de *b* es sólo una parte del contenido de *a*, tenemos $Ct\,(a) > Ct\,(b)$, de modo que la corroborabilidad de *a* será mayor que la de *b*. En vista de 8), *supra*, esto puede escribirse también $C\,(a,a) > C\,(b,b) = C\,(b,bc)$ o, como $a = bc$, tenemos

$$C\,(a,bc) > C\,(b,bc)$$

Este resultado puede formularse como sigue:

v) Si *b* se sigue de *a* pero no viceversa, y $0 \neq p\,(a) < p\,(b) \neq 1$, entonces existe un enunciado *d* (en nuestro caso, *d* puede ser, por ejemplo, la conjunción *bc*, es decir, *a* misma) tal que

$$C\,(a,d) > C\,(b,d)$$

Esta regla es válida para el grado de corroboración. *Pero la correspondiente regla para las probabilidades es falsa.* No sólo esto, sino que la negación exacta de esta regla es válida. Porque la siguiente regla es un teorema del cálculo de probabilidades.

$v)_p$ Si *b* se sigue de *a* pero no viceversa, y $0 \neq p\,(a) < p\,(b) \neq 1$ entonces, para todo enunciado *d*,

$$p\,(a,d) \leq p\,(b,d)$$

El choque entre grados de corroborabilidad y probabilidad no podría expresarse de forma más marcada que con el conflicto entre v) y $v)_p$. Pero v) se sigue directamente de nuestra idea fundamental de que el grado de corroboración es un informe sobre la severidad de las contrastaciones y que contrastabilidad es igual a la corroborabilidad. (Nada cambiaría aquí si volviéramos a decir «proporcional a» o incluso sencillamente «aumenta con» en lugar de «igual a».)

15) La situación aquí descrita es una consecuencia formal de la idea misma de que el grado de corroboración es un informe sobre contrastaciones y sobre su severidad. Con todo, v) reenuncia las observaciones intuitivas que se hicieron en 1) en la sección 27 respecto a las teorías de Maxwell y Fresnel; porque podemos considerar ahora que *a* es la teoría de Maxwell, *b* la de Fresnel, y *d* una contrastación (tal como la de Hertz) de la parte electromagnética de la teoría de Maxwell o la conjunción de una contrastación de este tipo y de una que sea

aplicable a ambas teorías (tal como la doble refracción, o el experimento de Fizeau).

16) Las condiciones expuestas hasta ahora —excepto las dos convenciones mencionadas bajo 3) y 6)— se derivan, todas, de nuestro principio fundamental de que el grado de corroboración es un informe sobre la severidad de las contrastaciones que ha superado una teoría. Pero ahora añadimos una idea nueva.

Tomemos, como antes, una hipótesis a de la cual sea deducible una hipótesis más débil b. (De nuevo podemos considerar que a es la teoría de Maxwell y b la de Fresnel.) Ahora suponemos que hemos contrastado b por medio de la contrastación x (el experimiento de Fizeau o la doble refracción) pero no la parte de a que va más allá de b. En este caso diremos que $C\ (a,x) < C\ (b,x)$, a pesar de la mayor corroborabilidad de a. Porque a contiene una parte, aún no contrastada, que puede ser refutada en la primera contrastación que intentemos. Pero si x contrasta también esta última parte (como lo hace el experimento de Hertz) entonces tenemos que $C\ (a,x) > C\ (b,x)$, siempre que la contrastación haya sido superada con éxito y suficientemente severa.

Podemos expresar esta consideración de otras varias formas. Por ejemplo, podemos decir que si b está bien corroborada por x y le añadimos a b una c completamente nueva que no esté en absoluto contrastada por x, entonces al $a = bc$ resultante estará peor contrastado por x que b, y, por tanto, menos bien corroborado, aunque su corroboración se elevará por encima de la de b tan pronto como se someta a unas cuantas contrastaciones severas a c, es decir, a la parte de a que va más allá de h. Podemos expresar todo esto mediante la exigencia:

Si x se explica por b sola, como mínimo en el mismo grado en que se explica por $a = b,c$, entonces $C\ (a,x) < C\ (b,x)$. Aquí «se explica» puede sustituirse por «se hace probable». Llegamos así a la formulación

vi) Si b se sigue de a; si $0 \leq p\ (a) < p\ (b) < 1$; y si $p\ (x,b) \geq p\ (x,a)$, entonces

$$C\ (a,x) < C\ (b,x)$$

Está claro que esta última exigencia es, como iii), una de ésas en las que el grado de corroboración se parece a la probabilidad. Porque la siguiente fórmula análoga es derivable del cálculo de probabilidades.

vi)$_p$ Si b se sigue de a; si $0 \leq p\ (a) < p\ (b) < 1$; y si $p(x,b) \geq p(x,a)$, entonces

$$p\ (x,a) < p\ (x,b)$$

En vista de esta nueva exigencia, la solución de nuestros dos problemas (de la sección 30) se hace aún más urgente.

Estas son todas las exigencias que tenemos que tener en cuenta. Primero encontré una definición de $C\,(a,b)$ que las satisface todas con ayuda de una definición sumamente intuitiva del *poder explicativo* de a respecto a b (véase la nota 21 en esta misma sección). Más tarde, encontré una definición más sencilla aún, como sigue (véase la definición D en la sección 31).

vii) Sea a una hipótesis coherente[24] y sea b la descripción del resultado de todos nuestros intentos de refutar a, de modo que $p\,(b) \neq 0$. Entonces

$$C\,(a,b) = \frac{p\,(a,b) - p\,(b)}{p\,(b,a) - p\,(a,b) + p\,(b)}$$

Con ayuda de la sola suposición de que p satisface el cálculo de probabilidades, puede mostrarse que todas nuestras exigencias son satisfechas; que la definición y, por tanto, el conjunto de exigencias, es coherente; y que $C\,(a,b)$ no es lo mismo que $p\,(a,b)$. Porque $C\,(a,b)$ no satisface, como $p\,(a,b)$ las reglas del cálculo de probabilidades.

De este modo se resuelven ambos problemas. Sin embargo, parece que no sólo hay cabida, sino necesidad real de dos mejoras más de nuestra definición vii).

En primer lugar, podemos generalizar nuestra definición introduciendo una tercera variable, digamos que c, que denote nuestra información básica[25]. (Esta, naturalmente, se representaba en la sección 31 por b.) Así llegamos a

viii) Sea a una hipótesis coherente[26]; sea b el resultado de auténticos intentos de refutar a; y sea c nuestra información básica (coherente)[27] compuesta de teorías que no estén sometidas a contrastación y también de condiciones iniciales. Entonces

$$C\,(a,b,c) = \frac{p\,(b,ac) - p\,(b,c)}{p\,(b,ac) - p\,(ab,c) + p\,(b,c)}$$

---

[24] No nocesitamos exigir coherencia si estamos preparados para complicar un poco nuestra fórmula insertando «— $p(\bar{a}.a)$» en el numerador y el denominador, es decir, escribiendo $C(a,b) = p(b,a) - p(\bar{a}.a) - p(b) \,/\, p(b,a) - p(\bar{a}.a) - p(ab) + p(b)$. Como resultado, $C(a,b) = -1$ siempre que a sea incoherente. Una observación correspondiente es aplicable a la definición relativizada viii).

[25] Esto presupone la relativización de $Ct(a)$: pondremos

$$Ct(a,c) = C(a,a,c) = 1 - p(a,c).$$

[26] Véase la nota 24.

[27] Véase la nota 24.

Esto corresponde exactamente a la definición D) de la sección 31, a la que se llegó por un argumento algo diferente.

Puede hacerse una advertencia técnica respecto a estas dos definiciones vii) y viii). Están basadas en *p (b,a)* —llamada la verosimilitud de *a* con respecto a *b* por R. A. Fisher— en vez de estarlo en *p (a,b)*. Este es un punto importante, porque de ese modo conservan todo su significado incluso si *p (a)* = 0; un caso que, como mostré en las secciones 80 y 81 de *L. I. C.*, puede suponerse válido en general, si *a* es una teoría universal[28]. Porque una teoría universal puede considerarse, *en un universo infinito* (infinito respecto al espacio o al tiempo), como un producto infinito de enunciados singulares y, por tanto, «infinitamente improbable»; o, más exactamente, con una probabilidad cero. Esto no le causa dificultades a nuestra teoría debido al hecho de que he desarrollado una generalización del cálculo de probabilidades habitual que da sentido a *p (x,y)* incluso aunque *p (y)* = 0. (Véase *L. I. C.*, Apéndice *IV.) En consecuencia, las expresiones «*p (b,a)*» o «*p (ba,c)*» y «*p (b,c)*» que aparecen en el *definiens* de vii) y de viii) respectivamente, tendrán sentido incluso aunque *a* [y *c*)] sean leyes con una probabilidad cero. Si en el *definiens* de vii) operásemos con «*p (a,b)*» en vez de «*p (b,a)*», entonces para cualquier ley universal *a* deberíamos obtener *C (a,b)* = 0/0, ya que en ese caso, *p (a,b)* — *p (a)* sería igual a cero y lo mismo ocurriría con *p (a,b)* = *p (a)*[29]. Este punto tiene una importancia considerable

---

[28] Véase también *L. I. C.*, Apéndice *VII. Que todas las leyes universales tienen una probabilidad cero y, por tanto, un grado de confirmación cero (si se define el grado de confirmación como una probabilidad) es también un resultado de Carnap (*Logical Foundations of Probability*, 1950, pp. 570 s.). Por consiguiente, las leyes universales no son confirmables en el sentido de Carnap. Por esta razón, Carnap sustituye (como hago yo) la probabilidad o grado de confirmación por otra idea —aunque él lo hace sólo *ad hoc*— cuando se trata de leyes universales. Pero la nueva función, que él llama «confirmación de ejemplos [o casos] restringidos», lleva a contradicciones, como señalé en *Brit. Jour. Phil. Science* 7, 1956, pp. 252 s. De hecho, uno puede obtener, con ayuda de la fórmula 17) en p. 573 del libro de Carnap, «grados de confirmación de ejemplos restrigidos» enormemente diferentes, *basándose en la misma evidencia, para dos leyes que difieran solamente en sus formulaciones verbales.*

[29] *El hecho de que p(a,ba) — p(a)* sea igual a cero si *a* es universal, mientras que *p(b)* > 0, es de un interés considerable en relación con el concepto de Carnap de «*confirmado*» (*Logical Foundations*, pp. 463 s.) Porque Carnap define «*Co(a,b)*», es decir, «*a es confirmada por b*» sirviéndose de una expresión que es equivalente a «*p(a,b) — p(a)* > 0», en nuestra notación; es decir, equivalente a «*p(a,b)* > *p(a)*». En consecuencia, ningún enunciado evidencial *b* «confirma una ley universal *a* si *p(a)* = 0 para las leyes universales» (como, en efecto, ocurre en la teoría de Carnap). *Así pues, las leyes no pueden ser «confirmadas»* en el sentido de Carnap. Es más, *Co* contradice el concepto de Carnap del grado de confirmación —que es una probabilidad, *p(a,b)*, por ejemplo— en presencia de la tautología «si *x* no está caliente, entonces nunca

y puede, desde luego, generalizarse con la introducción de c, es decir, de «información básica» y extenderse así a viii).

Nuestra segunda mejora es ésta: podemos hacer un poco más para que sea explícito nuestro requisito informal, la condición de que b consista en auténticos intentos de falsar a a; como se anunció en la sección 31, podemos formalizarlo en parte exigiendo que nuestros enunciados de contrastaciones empíricas sean inesperados o improbables a la luz de nuestra información básica, es decir, que su probabilidad, dada la información básica, sea (considerablemente) menor que 1/2.

Me parece que no podemos evitar un elemento intuitivo en nuestra definición, consistente en la condición de que «b» en «$C\ (a,b)$» denote una descripción de los resultados de *contrastaciones auténticas*. Porque pretendemos que $C\ (a,b)$ denote algo como la aceptabilidad de una teoría, una evaluación útil, aplicable y convincente de la teoría. En otras palabras, estamos tratando de un formalismo *interpretado* más que de uno sin interpretar; y nuestro elemento informal o intuitivo sólo interviene como una advertencia. «Si desea usted interpretar $C\ (a,b)$ o $C\ (a,b,c)$ como el grado de corroboración (o de aceptabilidad) de una teoría a, entonces tiene usted que asegurarse de que b describe los resultados de contrastaciones auténticas»[30].

Para resumir, quiero señalar dos cosas. En primer lugar, nuestro concepto de grado de corroboración toma plena consideración (automáticamente, por así decirlo) del *peso de la evidencia* que tantos problemas causa en las teorías de la probabilidad[31]. En segundo lugar, nuestra definición de los resultados comparativos deseados, no sólo en el caso de las teorías de

---

se da el caso de que x esté más caliente que y o igual de caliente que y»; o, más bien «si x no es confirmada por z e y es confirmada por z, entonces nunca se da el caso de que x esté mejor confirmada por z que y o igual de bien confirmada por z que y». Porque esto puede expresarse por «Si no-$Co(x,z)$ y $Co(y,z)$, entonces no $p(x,z) \geq p(y,z)$»; de donde obtenemos al instante, haciendo las sustituciones correspondientes a la equivalencia de Co mencionada *supra:*

1) Si $p(x,z) \leq p(x)$ y $p(y,z) > p(y)$, entonces $p(x,z) < p(y,z)$
2) Si $p(x,z) = p(x)$ y $p(y,z) > p(y)$, entonces $p(x,z) < p(y,z)$ (por 1)
3) Si $p(x,z) = p(x)$ y $p(y,z) > p(y)$, entonces $p(x) < p(y,z)$ (por 2).

Pero esto es absurdo; porque podemos escoger $p(x)$ tan cerca de 1 como queramos, o igual a 1; y si escogemos entonces $p(y,z) = 1/2$, por ejemplo, y $p(y) = 1/4$, por ejemplo, obtenemos una contradicción; lo mismo ocurre, más en general, si escogemos $y = xz$, y suponemos que x y z son independientes. (Véase *Brit. Jour. Phil. Science* 7, 1956, pp. 254 s.; también mis observaciones introductorias al Apéndice *IX de *L. I. C.*)

[30] Debe recordarse que existe un caso paralelo en la teoría de la probabilidad. Si $p(a,b)$ ha de interpretarse como el grado de creencia racional en a entonces b debe denotar todo el conocimiento de la evidencia que posee el creyente.

[31] *Cf.* mi «Tercera Nota» en el Apéndice *IX a *L. I. C.*

Maxwell y de Fresnel, sino también en el caso de las teorías de Newton y Einstein. E incluso nos permite expresar plenamente tanto el carácter algo dudoso de los resultados del eclipse como el hecho de que esos resultados son más favorables a Einstein que a Newton. Porque si a_1 es la teoría de Newton y a_2 es la de Einstein y b las mediciones del eclipse, obtenemos, ciertamente, que

$$p\ (b,a_1c) < p\ (b,a_2c)$$

porque los b estan *mucho* más alejados de la predicción de Newton que de la de Einstein. Por esto vemos al instante que b corrobora a Einstein mucho mejor que a Newton, de acuerdo con el sentido común y con lo que está generalmente aceptado.

Esto no se debe a ninguna peculiar sutileza de nuestra teoría, sino al simple hecho de que nuestra definición satisface nuestros requisitos. Contiene una evaluación directa de la severidad de las contrastaciones que ha superado la hipótesis y de cómo las ha superado de bien, o cómo de mal ha facasado en ellas.

Repito que no creo que mi definición del grado de corroboración sea una contribución a la ciencia excepto, quizá, que pueda ser útil como evaluación de las contrastaciones estadísticas (*cf.* Apéndice *IX de *L. I. C.*). Tampoco creo que constituya una contribución positiva a la metodología ni a la filosofía, excepto en el sentido de que puede ayudar (o eso espero) a aclarar las grandes confusiones que han surgido del prejuicio de la inducción y del prejuicio de que en ciencia buscamos probabilidades altas —en el sentido del cálculo de probabilidades— en vez de un alto contenido y unas contrastaciones severas.

33. EL HUMANISMO, LA CIENCIA Y EL PREJUICIO INDUCTIVO

No existe la inducción probabilística. La experiencia humana, en la vida cotidiana, así como en la ciencia, se adquiere fundamentalmente por el mismo procedimiento: la libre, injustificada e injustificable invención de hipótesis o suposiciones o expectativas y su subsiguiente contrastación. Las contrastaciones no pueden hacer «probable» a la hipótesis. Sólo pueden corroborarla, y eso sólo porque el «grado de corroboración» no es más que una etiqueta pegada a un informe, o una evaluación de la severidad de las contrastaciones superadas por la hipótesis.

Pero casi todos los meses se publican nuevas teorías sobre la inducción. Porque la aserción de que la probabilidad de una

ley aumenta con el número de ejemplos observados tiene una considerable fuerza intuitiva. He tratado de explicar esta fuerza intuitiva señalando que no se han diferenciado debidamente la probabilidad y el grado de corroboración. Tanto si mi explicación es satisfactoria como si no lo es, la actual superabundancia de teorías insostenibles sobre la inducción debe ser altamente insatisfactoria incluso para un inductivista.

En vista de la situación, quiero dirigirme a los autores de cualquier teoría futura de la inducción que quieran pretender para sus teorías algo parecido a la categoría científica. Recordando el famoso llamamiento de Kant a aquellos que se interesaban por la metafísica, trataré ahora de *convencer a todos aquellos que consideran que merece la pena tomarse un interés activo en la teoría de la inducción: que es inevitable, y aún más, necesario, que desistan, por el momento, de cualquier nuevo esfuerzo; que consideren todo lo que se ha hecho como deshecho; y que planteen la cuestión, antes de nada, de si es en absoluto posible tal cosa como una teoría probabilística de la inducción*[32].

Para ser más específico, desafío a cualquiera que crea posible aumentar la probabilidad de una teoría por medio de algún procedimiento inductivo, a explicar cuatro cosas:

i) Por qué los científicos prefieren invariablemente una teoría altamente contrastable cuyo contenido va mucho más allá de toda la evidencia observada a una hipótesis *ad hoc* concebida para explicar precisamente esa evidencia, y poco más que ella, a pesar de que la última hipótesis tiene que ser siempre mucho más probable que la primera sobre la base de cualquier evidencia dada. ¿Cómo ha de combinarse la exigencia de que una teoría tenga un alto contenido informativo —para el conocimiento— con la exigencia de una alta probabilidad, lo que significa falta de contenido y falta de conocimiento?

ii) Cómo evitar obtener probabilidades iguales a 1 para todas las leyes universales aún no refutadas, considerando que todas se hallan ejemplificadas casi en todo lugar; porque tanto la ley «todos los cisnes son blancos», es decir, «no existe cisne que no sea blanco» como la ley «todos los cisnes son no blancos», es decir, «no existe cisne blanco» están ejemplificadas en todas las regiones en las que no hay cisnes, es decir, según nuestro conocimiento actual, en casi todo el universo.

iii) Cómo evitar, en un universo infinito (o en uno prácticamente infinito) obtener la probabilidad cero para todas las

[32] *Cf.* Kant, *Prolegomena to Any Future Metaphysics that will be Able to Claim Scientific Status*, 1783. Las palabras en cursiva son de Kant, excepto que el término de Kant «metafísica», se ha sustituido por «teoría de la inducción».

leyes universales, considerando que una ley universal sobre un universo infinito puede expresarse siempre como el producto infinito de enunciados singulares. (Por ejemplo, «todos los cisnes son blancos» puede expresarse diciendo «*todas las cosas* tienen la propiedad *P*» donde «tener la propiedad *P*» se define por medio de la frase «o bien ser blanco o bien no ser cisne».)

iv) ¿Cómo responden a mis objeciones, planteadas en *L. I. C.* Apéndice *IX.

Este es mi desafío.

El principal argumento contra la inducción es el mismo que contra el idealismo: la inducción es demasiado fácil. Pero en el caso de la inducción —a diferencia del caso del idealismo— puede utilizarse la argumentación para mostrar lo absurdo de la teoría. Porque desde Bacon, toda inducción consiste en la recolección y tabulación (estadística) de ejemplos, especialmente de ejemplos confirmadores.

Pero los ejemplos confirmadores, las «verificaciones», pueden conseguirse con sólo pedirlos. *De hecho, cualquier «ejemplo», con la sola excepción de un ejemplo falsador, confirma toda teoría universal.* De ahí que no merezca la pena tener ejemplos confirmadores. (Véase, por ejemplo, la sección 28 de *L. I. C.*, especialmente la nota *1, y la sección 80, especialmente la nota *4.)

El que los ejemplos sean tan fáciles está estrechamente relacionado con la ley de la exclusión del término medio. Esto puede verse fácilmente si expresamos nuestras teorías en la forma de un enunciado de «no-existe», es decir, de la negación de un enunciado existencial (como puede hacerse con toda teoría universal; véase sección 15). El enunciado «no existe un cisne no blanco» resulta confirmado de hecho por cualquier cosa en el mundo, con la excepción de un cisne no blanco, ya que es equivalente al enunciado «todas las cosas tienen la propiedad de no ser un cisne no blanco». Del mismo modo, el enunciado «no existe una máquina en perpetuo movimiento» es equivalente al enunciado «todas las cosas tienen la propiedad de no ser máquinas en perpetuo movimiento». Por tanto, todas las cosas son ejemplos verificadores de este enunciado con la sola excepción de un ejemplo que sea, en efecto, una máquina en perpetuo movimiento[33]. Esta es la principal razón por la que

[33] La aserción que se hace aquí (*cf.* secciones 28 y 80 de *L. I. C.*) está estrechamente relacionada con lo que C. G. Hempel ha llamado la «paradoja de la confirmación». Hempel parece pensar (*Mind*, N.S. 55, 1946, p. 79) que las «paradojas» son sólo «peculiaridades» y «en un análisis más atento... demuestran ser razonables», de modo que debemos aceptar, efectivamente, el hecho de que un pañuelo blanco o un cuervo negro *confirman* el enunciado «todos los cisnes son blancos». Puede que sea así. Pero pienso que todos esos

sólo los auténticos intentos de refutar una teoría pueden contar.

«Nada hay tan cierto, ni siquiera en la ciencia matemática, como la incapacidad de una serie de hechos científicos por sí solos para llevar a un descubrimiento», escribió Brewster hace más de cien años, expresando, quizá, una actitud más madura que la nuestra, hacia la ciencia[34]. Porque hoy se nos dice que las leyes y las teorías universales no son realmente necesarias para la ciencia, que la ciencia puede inferir sus predicciones

---

ejemplos «confirmadores» o «verificadores» son claramente demasiado fáciles como para aceptarlos como *corroboraciones*, porque no son, en general, el resultado de auténticos intentos de refutar la teoría de que «todos los cisnes son blancos».

El concepto de Hempel sobre la «confirmación» o el «ejemplo confirmador» (*Journal of Symbolic Logic* 8, 1943, pp. 122-143) es verificacionista: la idea es que el enunciado x confirma el enunciado y si, y sólo si, x describe un *ejemplo* verificador de y. Pero debido a lo inadecuado de los «requisitos lógicos» de Hempel, su definición tiene las siguientes consecuencias, evidentemente no buscada: Sea a el nombre de una cosa individual. Entonces, el enunciado singular «a es un cisne blanco» no sólo confirma, como se esperaba, «todos los cisnes son blancos» y «todas las cosas no-blancas son no-cisnes», sino que también confirma (lo que no se buscaba en absoluto) «todas las cosas son blancas», «todas las cosas son cisnes», «todas las cosas son cisnes blancos»; y excepto por su propia exclusión *ad hoc* de la identidad (sobre esto, véase *infra*) «sólo hay un cisne blanco», «existe una, y sólo una, cosa, y ésta es un cisne blanco». Además, nuestro enunciado singular confirma *sólo los enunciados que sean compatibles con los enunciados aquí citados*, de modo que confirma «todos los cisnes son blancos» sólo en el sentido pickwickiano de que significa «todos los cisnes son blancos porque, aparte del a que hemos visto, no existe ninguno». Esto muestra que la definición de Hempel es, evidentemente, *demasiado estrecha y demasiado amplia*. Confirma enunciados que no se proponía confirmar y deja sin confirmar los que se proponía confirmar. Por ejemplo, los dos enunciados «todo ser humano tiene una madre humana» y «todo ser humano hembra tiene una madre humana» son ambos inconfirmables por seres humanos (aunque son confirmados por «a es no humano»), según la teoría de Hempel (puesto que no tienen un «desarrollo» finito, en el sentido de Hempel, entre seres humanos); pero el enunciado «todo ser humano varón tiene una madre humana» es confirmable, por ejemplo, por el enunciado «María es la madre de Eduardo» (y María es un ser humano hembra y Eduardo es un ser humano varón). Respecto a la exclusión de la identidad en el lenguaje modelo de Hempel, no pienso que pueda hacerse más adecuada una definición mediante el intento de impedirnos expresar en nuestro lenguaje modelo consecuencias que son palpablemente contra-intuitivas. (*Cf.* también J. W. N. Watkins, «Between Analytical and Empirical», *Philosophy*, 1957, pp. 116-123 y pp. 125-127.)

Por cierto, se me ha criticado por no dar una definición precisa ni de la «*ejemplificación de una ley*» ni de un «*ejemplo refutador de una ley*». Sin embargo, doy definiciones «semánticas» de esos conceptos en la última nota a pie de página de mi «Nota sobre la definición de verdad de Tarski», ahora en *Conocimiento Objetivo (Obj. Kn.)*, pp. 302-306.

[34] Sir David Brewster, *Memoirs of the Life, Writings and Discoveries of Sir Isaac Newton*, segunda edición, 1860, vol II, p. 328.

(con probabilidad) a partir de una mera colección de hechos singulares[35].

Dos actitudes o tendencias que se encuentran a veces juntas fomentan la creencia en la inducción. Una es el deseo de una autoridad suprahumana: la autoridad de la ciencia, muy por encima de los caprichos humanos y ejemplificada por la ciencia «exacta» de la matemática y por las ciencias naturales, en la medida en que se basan, firme y claramente, en hechos: hechos verificados, confirmados. La otra es el deseo de ver en la ciencia no la obra de una inspiración o revelación del espíritu humano, sino, más o menos, una recopilación mecánica que, en principio, puede ser realizada por máquinas. (Porque, ¿qué somos sino máquinas?) En el fondo, las dos tendencias pueden ser una: la tendencia a desbancar al hombre.

Ahora bien, un poco de desbancamiento puede hacernos mucho bien, sobre todo si se hace con la gracia y el buen humor de Bertran Russell. «Puesto en una cáscara de nuez» —escribió en su prefacio a *Mysticism and Logic*— «el cambio en mi perspectiva se reduce a esto: que ya no considero la solemnidad como un medio de alcanzar la verdad; la observación de la vida muestra que la gente solemne es, por lo general, farsante y que los estados de ánimo solemnes también tienen algo de farsa».

Pero hay otros desbancadores, y, entre ellos, algunos bastante solemnes. Y, al contrario de quienes honran a la Ciencia con mayúscula, porque está verificada, o (como he desbancado a la verificación) «está confirmada» o «es exacta», hay no pocos que creen que aunque se hace «firme» por medio de la confirmación[36], no es profunda; que, puesto que el mundo no tiene profundidad, la ciencia no necesita teorías, sino sólo colecciones de hechos cuya frecuencia puede inducir la probabilidad de sus repeticiones futuras.

Yo veo la ciencia de un modo muy diferente. Respecto a su autoridad, o confirmación o probabilidad, creo que es nula; es todo conjetura, *doxa* más que *epistèmè*. Y la teoría de la probabilidad incluso me «confirma» en esto, atribuyendo una probabilidad cero a las teorías universales.

Pero vista como el resultado del esfuerzo humano, de los sueños, esperanzas, pasiones humanas y, sobre todo, como el resultado de la admirable unión de la imaginación creadora y del pensamiento racional y crítico, me gustaría escribir Ciencia, con la mayúscula más grande que pueda encontrarse en la caja alta del impresor.

---

[35] Véase el texto a las notas 110 y 111 a la sección 13, capítulo I *supra*.
[36] *Cf. Philosophy of Science* 12, 1945, nota 2 en p. 99.

La ciencia no es sólo, como el arte y la literatura, una aventura del espíritu humano, sino que entre las artes creativas es quizá la más humana: llena de fallos e imprevisiones humanas, muestra esos fogonazos de discernimiento que nos abren los ojos a las maravillas del mundo y del espíritu humano. Pero esto no es todo. La ciencia es el resultado directo del más humano de los esfuerzos humanos, el de liberarnos. Forma parte de nuestro esfuerzo por ver más claro, por entender el mundo y entendernos a nosotros mismos, por actuar como seres adultos, responsables e ilustrados: «La ilustración», escribe Kant, «es la emancipación del hombre del estado de tutela auto-impuesta... de un estado de incapacidad para usar su propia inteligencia sin dirección externa. A ese estado de tutela le llamo 'auto-impuesto' si no se debe a ninguna falta de inteligencia, sino a la falta de valor y determinación para usar la propia inteligencia en lugar de confiar en un jefe. *¡Sapere aude!* ¡Atreveos a utilizar vuestra inteligencia! Esta es la máxima de la ilustración»[37].

Kant nos reta a utilizar nuestra inteligencia en lugar de confiar en un jefe, en una autoridad. Esto debería tomarse como un reto a rechazar incluso al experto científico como jefe, o incluso a *la propia ciencia.* La ciencia no tiene autoridad. No es el producto mágico de lo dado, los datos, las observaciones. No es un evangelio de verdad. Es el resultado de nuestros propios esfuerzos y errores. Somos usted y yo los que hacemos la ciencia lo mejor que podemos. Somos usted y yo los que somos responsables de ella.

La ciencia, uno puede sentirse a veces tentado a decir, no es más que sentido común ilustrado por el pensamiento crítico e imaginativo. Pero es más. Representa nuestro deseo de saber, nuestra esperanza de emanciparnos de la ignorancia y de la estrechez de miras, del miedo y la superstición. Y esto incluye la ignoracia del experto, la estrechez de miras del especialista, el miedo a que se demuestre su equivocación, su «inexactitud» o el miedo a fracasar en el intento de demostrar o justificar nuestro caso. E incluye la creencia supersticiosa en la autoridad de la ciencia misma (o en la autoridad de los «procedimientos inductivos» o de la «pericia»).

La bomba atómica (y posiblemente también el llamado «uso pacífico de la energía nuclear» cuyas consecuencias pueden ser incluso peores a largo plazo) nos ha mostrado, pienso yo, la superficialidad del culto a la ciencia como un «instrumento» de nuestro «dominio de la naturaleza»: nos ha mostrado que este dominio, este control, es capaz de autodestruirse y es más

[37] I. Kant, *Was ist Aufklärung?* (¿Qué es ilustración?), 1785.

capaz de esclavizarnos que de hacernos libres, si no nos elimina por completo. Y, aunque merece la pena morir por el conocimiento, no merece la pena morir por el poder. (El conocimiento es una de las pocas cosas por las que merece la pena dar la vida, junto con la libertad, el amor, la bondad y la ayuda a quienes la necesitan.)

Todo esto puede estar muy trillado. Pero hay que decirlo de vez en cuando. La Primera Guerra Mundial no sólo destruyó la república del saber; estuvo a punto de destruir la ciencia y la tradición del racionalismo, porque hizo a la ciencia técnica, instrumental. Llevó a una especialización creciente y apartó de la ciencia a quienes tendrían que ser sus verdaderos usuarios: el aficionado, el amante de la sabiduría, el ciudadano corriente, responsable, que tiene un deseo de saber. Todo esto empeoró mucho con la Segunda Guerra Mundial y la bomba. Por eso hay que volver a decir estas cosas. Porque nuestras democracias atlánticas no pueden vivir sin ciencia. Su principio fundamental —aparte de ayudar a reducir el sufrimiento— es la verdad. No pueden vivir si dejamos que se destruya la tradición del racionalismo. Pero lo que podemos aprender de la ciencia es que es difícil llegar a la verdad: que es el resultado de derrotas indecibles, de intentos descorazonadores, de noches en vela. Este es uno de los grandes mensajes de la ciencia y no creo que podamos pasarnos de él.

Pero es precisamente este mensaje el que la especialización moderna y la investigación organizada amenazan con socavar; y la resurrección de las ingenuas concepciones de Bacon sobre la inducción —de la creencia de que la ciencia consiste en la recolección y tabulación de ejemplos, y sobre todo de ejemplos confirmadores— debe ser combatida por los que creen en la razón humana.

Tal como recuerdo ahora mi experiencia en la Sociedad Aristotélica (que relaté en la sección 1), veo ahora que mi intento de explicar mis concepciones no podía dejar de fracasar. Había demasiado en él y demasiado contra él. Puede que, ahora ya, haya mostrado la insuficiencia del precioso método de enunciar el caso de uno con claridad y sencillez, y dejarlo ahí, sin discutir todos esos «ismos» que pueden suavizar el impacto de cualquier idea nueva. Yo intenté efectivamente ese método: enuncié mi solución del problema total y sencillamente en las secciones 2 y 3. En cierto modo, no hubiera hecho falta nada más que eso, y es posible que algunos lectores hayan sentido toda la fuerza del argumento. Pero no creo que el largo camino que hemos recorrido desde entonces haya sido innecesario. Si acaso, dudo si me habré adentrado lo suficiente en todas las suposiciones y los prejuicios que acosan a este pro-

blema lógico tan sencillo y tan directo y que obstaculizan el examen imparcial de cualquier solución sencilla y directa.

Desde luego, mientras escribo esto, me siento inclinado a pensar que esta vez he conseguido hacer comprender mi punto de vista. Pero cuando trato de adoptar una perspectiva un poco más apartada del asunto, lo pongo en duda. Las personas no sólo oyen lo que esperan oír, sino que no leen más que lo que quieren leer. Y como siempre digo que sólo podemos captar las cosas con ayuda de nuestras teorías, seré el último en quejarme de ello. Sin duda hay dos formas posibles de interpretar las cosas mientras las leemos: un lector puede ser acrítico o crítico hacia sus propias interpretaciones. Pero me temo que un tiempo y un lugar en los que los lectores intenten habitualmente refutar sus propias interpretaciones y expectativas sobre lo que están leyendo sólo existen en los sueños del escritor.

Addendum: Observaciones críticas sobre el análisis de significados

La razón por la que no quiero presentar mi definición del grado de corroboración como meta alcanzada(véase la sección 30) es que —al contrario, me temo, que la mayoría de los filósofos— estoy convencido de la verdad de las tres proposiciones siguientes.

1) *Las preguntas de qué-es,* tales como *¿qué es la Justicia?* o *¿qué es el grado de corroboración?* son siempre inútiles, carecen de interés filosófico o científico; y lo mismo ocurre con todas *las respuestas* a las preguntas de qué es, tales como las *definiciones*. Hay que admitir que algunas definiciones pueden ser útiles a veces para contestar a otras preguntas: cuestiones urgentes que no pueden descartarse; dificultades auténticas que pueden haberse planteado en la ciencia o en la filosofía. Pero las preguntas de qué es, en cuanto tales, no plantean ese tipo de dificultad.

2) Da lo mismo que una pregunta de qué-es se plantee para inquirir sobre la esencia o la naturaleza de una cosa o que se plantee para inquirir sobre el significado esencial o el uso apropiado de una expresión. Esas clases de preguntas de qué-es son fundamentalmente iguales. Una vez más, hay que admitir que la respuesta a una pregunta de qué-es —por ejemplo, una respuesta que señale distinciones entre dos significados de una palabra que se hayan confundido con frecuencia— puede no carecer de utilidad, siempre que la confusión haya llevado a dificultades serias. Pero en ese caso, no es la cuestión de qué-es lo que estamos tratando de resolver, son ciertas contradicciones que surgen de nuestra dependencia de ideas algo ingenuas. (El

primer ejemplo que se trata más adelante —el de las ideas de una derivada y de una integral— proporcionará una ilustración de este caso.) La solución puede muy bien ser la *eliminación* (en vez de la clarificación) de la idea ingenua, y su sustitución por una totalmente diferente[38]. Pero una respuesta a una pregunta esencialista (la pregunta de qué-es) nunca es fructífera. Porque las palabras, o los conceptos o las nociones, nunca son nada más que meros instrumentos, útiles para formular nuestras teorías. (A pesar de lo contrario que soy a la interpretación instrumentalista de las *teorías* científicas, soy instrumentalista en lo que respecta a las *palabras* o *conceptos* o *nociones*)[39].

En consecuencia, no puede ser el objetivo principal de la filosofía (ni de ninguna empresa racional ni crítica) el clarificar o definir ideas o conceptos o nociones o significados, ni sustituir algunas ideas o conceptos o nociones o significados dados por otros más exactos.

3) El problema, en especial, de sustituir un término «inexacto» —por ejemplo, el problema de dar una definición en términos «exactos» y «precisos»— es un pseudo-problema. Depende, esencialmente, de los inexactos e imprecisos términos «exacto» y «preciso». Estos son muy engañosos, no sólo porque sugieren con fuerza que existe lo que no existe —la exactitud o precisión absoluta—, sino también porque tienen una gran carga emocional: Bajo la apariencia de carácter científico y objetividad científica, sugieren que la precisión o exactitud es algo superior, una especie de principio último, y que está mal, es acientífico o denota atontamiento, el usar términos inexactos (como está mal, de verdad, no hablar tan clara y sencillamente como sea posible). Pero no hay tal cosa como un término «exacto» o términos hechos «precisos» por medio de «definiciones precisas». Además, una definición debe usar siempre términos no definidos en su *definiens* (puesto que, de otro modo, nos veríamos envueltos en una retrogresión infinita o en un círculo); y si tenemos que operar con una serie de términos no definidos, no importa mucho si usamos unos pocos más. Naturalmente, si una definición contribuye a resolver un problema auténtico, la situación es diferente; y algunos problemas no pueden resolverse sin aumento de la precisión. En realidad, ésta es la única forma en la que podemos hablar razonable-

[38] [*Cf.* la discusión de «diálisis» en *Búsqueda sin término*, sección 7. *Ed.*]

[39] He descrito esta posición, en varios lugares, como nominalismo metodológico en oposición tanto al esencialismo metodológico, que plantea y responde preguntas de *qué-es* como al nominalismo metafísico que afirma que las palabras con significado no son *más* que nombres de cosas (o imágenes de cosas en la memoria) y que, por tanto, da una respuesta esencialista a una pregunta (implícita) esencialista y una respuesta errónea, para colmo.

mente de precisión: la exigencia de precisión es vacía a menos que se plantee *en relación* con ciertos requisitos que surgen de nuestros intentos de resolver un problema determinado.

En esta sección voy a defender estas tres proposiciones.

Van en sentido contrario, mucho me temo, a la principal corriente o tendencia o rumbo de la filosofía contemporánea —incluida la filosofía de la ciencia— que tiende a combinar una interpretación instrumentalista de las *teorías* con una interpretación esencialista de los *conceptos*. Esta tendencia está ilustrada por un pasaje característico de Schlick en el que expone ciertas ideas de Wittgenstein. Schlick nos dice que las famosas preguntas de qué-es en los diálogos de Platón, tales como la pregunta «¿qué es la Justicia?» hecha por el «Sócrates» de *La República*, muestra que la filosofía de Sócrates estaba dedicada a «lo que podemos llamar 'la búsqueda del significado'. El trataba de clarificar», escribe Schlick, «el significado de nuestras expresiones y el sentido real de nuestras proposiciones... Diré breve y claramente», Schlick resume su posición, «que creo que la ciencia debe definirse como *la búsqueda de la verdad* y la filosofía como *la búsqueda del significado*»[40].

La definición de Schlick de la naturaleza esencial de la filosofía sigue siendo todavía muy influyente. Resultará obvio, por mis dos primeras tesis, en esta sección, que yo no pienso que tal definición pueda decirnos nada que merezca la pena (las definiciones nunca comunican un significado, excepto a los que necesitan un diccionario) o que se acerque lo más mínimo a una clarificación de la tarea y objetivo de la mayoría de los filósofos (quienes, si merecen ser escuchados, son buscadores de

[40] M. Schlick, «The Future of Philosophy» (*Publications in Philosophy*, editado por el College of the Pacific, 1932), citado aquí de Schlick, *Gesammelte Aufsätze* 1926-1936 (publicado en 1938), p. 126. Schlick dice que al científico *qua* científico, sólo le interesa la verdad —la verificación— y no el significado. Pero —Schlick pregunta— ¿cómo puede un científico contrastar la verdad de un enunciado cuyo significado no entiende? Sobre todo, si el significado del enunciado es el método de su verificación. Y en p. 127, Schlick escribe: «En la medida en que el científico descubre el significado oculto de las proposiciones que utiliza en su ciencia, es un filósofo». Y menciona como ejemplo el descubrimiento de Newton del concepto de masa y «el análisis de Einstein del significado de la palabra 'simultaneidad' tal como se usa en física» Discutí y descarté el último ejemplo en *La sociedad abierta (O. S.)*, capítulo II, sección II. El primero, el concepto de masa de Newton (que Newton definió como la integral espacial de la densidad) es un ejemplo desafortunado que hubiera resultado muy embarazoso para Newton; de hecho, el ejemplo es incomprensible excepto se supone que en lo que Schlick estaba pensando no era en la definición de Newton de «masa», sino en la idea de materia, junto con la interpretación esencialista de la teoría de Newton sobre la gravedad, que el propio Newton rechazó. He discutido esta interpretación en «Tres concepciones sobre el conocimiento humano», sección III, reimpreso en *El desarrollo del conocimiento científico (C. & R.)*.

la verdad). Por cierto no creo que Schlick represente correctamente la concepción del Sócrates histórico. Pero estoy de acuerdo en que Platón (y el «Sócrates» platónico de *La República*)[41] estaba muy preocupado por cuestiones de qué-es. De hecho, la creencia de que la filosofía debe analizar el significado de las palabras o los conceptos, que tiene que responder a preguntas de qué-es, de que debe dar definiciones, deriva de las metafísicas de Platón y Aristóteles. Pero se remonta a raíces más profundas aún: al animismo. (Esta puede ser una de las razones por las que es tan difícil de erradicar)[42]. Como lo muestra nuestro pasaje de Schlick, esta creencia no fue superada por los oponentes nominalistas y positivistas del idealismo platónico y está presente en Berkeley, Wittgenstein y sus seguidores, tanto como en el neo-aristotelismo. Dicha creencia inspira lo que Edmund Husserl ha llamado «la fenomenología pura» o «la intuición de las esencias»; lo que G. E. Moore llama «análisis filosófico» y lo que Rudolf Carnap llama «la explicación de los conceptos»[43]. Por profundas que sean las

[41] Para una discusión del problema socrático (es decir, el problema de la relación entre el platonismo y el Sócrates histórico), véase *La sociedad abierta (O. S.)*, nota 56 al capítulo 10, y Richard Robinson en *The Philosophical Review* 60, 1951, especialmente pp. 494 s.

[42] La creencia animista en el *poder de las palabras* —en la magia de las palabras—, está, conjeturo yo, implicada en el proceso por el que el niño aprende a hablar; un proceso ligado a su experiencia de controlar su entorno por medio de ruidos pre-lingüísticos. Pienso que esta creencia está implicada en lo que queremos decir cuando hablamos del «significado de las palabras»; lo que quizá sea una de las razones por las que algunas personas creen que el análisis de significados es importante. Hay que agradecer a Freud su profunda comprensión del animismo y de la creencia en la «omnipotencia de los pensamientos», símbolos y palabras. Sin embargo, Freud nunca sospechó que su propio enfoque esencialista tenía una base animista. Esto resulta claro en uno de sus últimos artículos (*Collected Papers* 5, 1952). En él habla del problema de «acercarse más a la *naturaleza* o, como dice la gente a veces, a la *esencia* de lo mental» (p. 377). Y declara que «ser consciente no puede ser la esencia de lo que es mental», y aún más (con un agradecimiento a T. Lipps) «que lo mental es en sí mismo inconsciente y que lo inconsciente es lo verdaderamente mental» (p. 382). Quizá el punto más interesante en esta discusión algo inútil (inútil porque preguntar por la esencia de lo mental es como preguntar por el espíritu [alcohol] del coñac o por el ánima del alma; y si lo inconsciente es la esencia de lo mental, entonces la energía mental será la esencia de lo inconsciente y la libido la esencia de la energía mental...) sea la percepción de Freud (p. 378) de que un físico *no* preguntaría «¿cuál es la naturaleza de la electricidad?» Esto, según Freud, plantea una objeción porque «la psicología también es una ciencia natural». El rechazo de Freud de esta objeción no se deduce de sus argumentos, algo ambivalentes. Y cuando sugiere que en psicología «el caso es diferente» porque todo el mundo se «comporta como... un psicólogo aficionado» y todo psicólogo aficionado cree en las esencias, se acerca mucho (pero no lo suficiente) a decir que el esencialismo es una actitud pre-científica, y a descubrir así, sus raíces mágicas y animistas.

[43] Hume habla de la «explicación exacta» de palabras (como «poder», *cf.*

diferencias entre estas escuelas, todas ellas consideran que la tarea de la filosofía es clarificar el significado de las ideas o de los conceptos o de las palabras; y por esta razón puede tratarse de todos ellos juntos aquí.

No quiero repetir aquí lo que he dicho en otros lugares sobre las esencias, del análisis de significados y las definiciones[44]. En cambio intentaré defender las tres tesis formuladas al principio de esta sección analizando una serie de contra-ejemplos *prima facie*; es decir, una serie de casos que, puede afirmarse, refutan mis tres tesis, ya que establecen la importancia de las definiciones que clarifican, explican o precisan el significado de algunos conceptos importantes. Mi primer ejemplo será la definición de la derivada (el cociente diferencial o medida de la inclinación de una curva) como límite de un cociente de diferencias. Este ejemplo se discutirá con cierto detalle. Mi segundo ejemplo será la «teoría de las descripciones» de Russell, una especie de ejemplo típico en este campo. Mis ejemplos tercero y cuarto serán la definición de la dimensión según la teoría de conjuntos y la definición de la verdad de Tarski.

En cada uno de estos casos supondré, sin mayor discusión, que los ejemplos pueden ser considerados por los devotos del análisis o la explicación de significados como refutaciones de mis tres tesis. Mi objetivo será mostrar, por el contrario, que las apoyan: que no son ejemplos de «clarificación» o «explicación» de significados, sino algo diferente: intentos de resolver problemas concretos, tales como la eliminación de contradicciones; y que logran su objetivo (al menos en los dos primeros casos) mediante el *abandono* del intento de clarificar o hacer «exacto» o «explicar» el significado deseado o intuitivo de los conceptos en cuestión.

Mi primer ejemplo es muy instructivo. Se ha afirmado que

---

su nota a pie de página en la parte II de la sección IV de *Enquiry*) con lo que quiere decir «fijar el significado preciso» de una palabra y «suprimir con ello parte de esa oscuridad de la que tantas quejas hay en... la filosofía» (*op. cit.* sección VII).

Carnap, en el primer capítulo «On Explication», de su *Logical Foundations of Probability*, 1950, se refiere a Kant, Husserl, Moore y Langford. Supone, explícitamente, que las preguntas de qué-es son problemas pertinentes; y sólo critica a quienes «empiezan a buscar una respuesta» —es decir, a una pregunta de qué-es— sin asegurarse antes de que «los términos de la pregunta» son suficientemente claros o exactos (p. 4). La tarea de «*explicar*» un «*explicandum*» se define como la de transformar el *explicandum* en (o sustituirlo por) un concepto exacto, el *explicatum*.

[44] Véase la nota 2 a la sección 19 de *L. I. C.* y la sección II del capítulo 2 de *La sociedad abierta (O. S.)* (véase también la sección 10 de *La miseria del historicismo* [*The Poverty*]). [*Cf. Búsqueda sin término (U. Q.)*, sección 7. *Ed.*]

la historia moderna del problema de los fundamentos de las matemáticas es, en gran medida, la historia de la «clarificación» de las ideas fundamentales del cálculo diferencial o integral. El concepto de *derivada* (la inclinación de una curva o tasa de incremento de una función) se ha hecho más «exacta» o «precisa» definiéndola como el límite del cociente de diferencias (dada una función derivable); y el concepto de una integral (el área o «cuadratura» de una región cerrada por una curva) ha sido también definida con «exactitud». El ejemplo tiene una importancia crucial para nuestro problema. Los intentos de eliminar las contradicciones en este terreno constituyen no sólo uno de los motivos principales del desarrollo de las matemáticas durante los últimos cien o incluso doscientos años, sino que también han motivado la investigación moderna de los «fundamentos» de las diversas ciencias y, en particular, la moderna búsqueda de precisión y exactitud. «Así, los matemáticos», dice Bertrand Russel, escribiendo sobre una de las fases más importantes de esta evolución, «sólo despertaron de sus 'sopores dogmáticos' cuando Weierstrass y sus seguidores mostraron que muchas de sus más caras proposiciones son, en general, falsas. Macaulay, comparando la certeza de las matemáticas con la incertidumbre de la filosofía, pregunta que quién ha oído jamás hablar de una reacción contra el teorema de Taylor. Si hubiese vivido ahora, él mismo podría haber oído hablar de tal reacción, porque ése es precisamente uno de los teoremas que han derrocado las modernas investigaciones. Golpes tan duros como ése a la fe matemática han producido el amor por el formalismo que les parece, a los que ignoran los motivos, simplemente una pedantería atroz»[45].

Quizá sea excesivo leer en este pasaje de Russel su aceptación de una concepción que yo considero verdadera: que sin «golpes tan duros como ése» —es decir, sin la urgente necesidad de eliminar las contradicciones— el amor por el formalismo es, en efecto, «simplemente una pedantería atroz». Pero pienso que Russell expresa efectivamente su concepción de que sin una necesidad urgente, un problema urgente que resolver, la mera exigencia de precisión es indefendible.

Pero esto no es más que una cuestión mínima. La principal cuestión es ésta: la mayoría de la gente, incluidos los matemáticos, consideran la definición de la derivada, en términos de límites de sucesiones, como si fuera de una definición en el sentido de que analiza, o precisa o «explica», *el significado* intui-

[45] Bertrand Russell, *Mysticism and Logic*, 1918; *cf.* el penúltimo párrafo del capítulo V.

tivo del *definiendum* de la derivada. Pero esta creencia tan extendida es errónea.

Podemos ilustrar dos sentidos (entre muchos) de la palabra «definición» con unas pocas definiciones del número 23.

Podemos escribir:

a)	$x = 23$ si, y sólo si,	$x = 20 + 3$
b)		$x = 2 \times 10 + 3$
c)		$x = 2^4 + 2^2 + 2^1 + 2^0$
d)		$x = 3^3 - 2^2$
e)		$x = 2.018 - (15 \times 133)$

La mayor parte de la gente dirá que las equivalencias *a)* y *b)* pueden definir «23» y quizá también *c)*, si queremos una definición en términos de potencias de 2. Pero ni *d)* ni *e)* serán una «definición razonable» de «23», aunque son, desde luego, verdaderas equivalencias; porque de ningún modo analizan ni dilucidan ni explican el significado de «23». Nadie que entienda el significado de «23» lo esclarecería o «explicaría» por medio de *d)* o *e)*.

Ahora bien, lo que quiero argüir es esto: No sólo *d)* y *e)* son definiciones perfectamente legítimas en el sentido general del término «definición» que es ahora aceptado por la mayoría de los matemáticos y los lógicos[46], sino que *d)* y *e)* pueden servir muy bien para ilustrar justamente ese sentido de la palabra «definido» en el que la derivada es definida como un límite del cociente de diferencias; un sentido que, mucho me temo, resulta oscurecido en la forma habitual de tratar el tema.

Newton y Leibniz y sus sucesores no negaron que una derivada o una integral pudiera *calcularse* como el límite a ciertas sucesiones, aproximadamente en el sentido en el que *d)* o *e)* pueden usarse para calcular x. Pero no hubieran considerado esos límites como posibles definiciones porque no dan el *significado,* la *idea,* de una derivada o de una integral[47].

Porque una derivada es la medida de una velocidad, de la inclinación de una curva. Ahora bien, la velocidad de un cuerpo en un determinado instante es algo real, un atributo concreto (relacional) de ese cuerpo en ese instante. En contraposición, el límite de una sucesión de velocidades medias es algo sumamente abstracto, algo que existe solamente en nuestros pensamientos. las velocidades *medias* son, ellas mismas, irreales. Su sucesión interminable lo es más aún; y el límite de esta sucesión sin fin es una construcción puramente matemática a partir de esas entidades irreales. Ahora bien, es intuitiva-

[46] Véase A. Tarski, *Logic, Semantics, Metamathematics,* 1956, p. 299.
[47] *Cf.* R. Courant y H. Robbins, *What is Mathematics?,* 1941, p. 433.

mente bastante obvio que este límite debe coincidir numéricamente con la velocidad y que, si el límite puede calcularse, podemos calcular a partir de él la velocidad. pero según las concepciones de Newton y de sus contemporáneos, sería empezar la casa por el tejado, si *definiésemos* la velocidad como idéntica a su límite, en vez de como un estado real del cuerpo —en un instante determinado o en un punto determinado de su trayectoria— a calcular por medio de cualquier invención matemática que seamos capaces de imaginar.

Lo mismo ocurre, naturalmente, con la inclinación de una curva en un punto dado. Su medida será igual al límite de una sucesión de medidas de otras determinadas inclinaciones medias (en vez de inclinaciones reales) de esta curva. Pero no es, en su significado o esencias propios, un límite de una sucesión: la inclinación es algo que podemos dibujar a veces realmente en un papel, y construir con reglas y compases, mientras que un límite es, en esencia, algo abstracto, que rara vez se alcanza o realiza, a lo que sólo puede aproximarse, cada vez más, una sucesión de números.

Del mismo modo, una sucesión de polígonos inscritos puede aproximarse a un círculo, y podemos *calcular* su área o «cuadratura» de esta manera, *aproximadamente* (con cualquier grado de aproximación que se desee); pero sería puro absurdo, desde el punto de vista de la comprensión de su significado o esencia, decir que un círculo *es* el límite (nunca realizado) de esta sucesión de polígonos inscritos. Porque es el círculo el que define esta sucesión —como la sucesión de polígonos inscritos en este círculo— y no al revés. Del mismo modo, está claro que el área del círculo será igual al límite de las áreas de los polígonos inscritos; pero sería puro absurdo tratar de explicarlo de ese modo; igual que sería absurdo tratar de describir el cuerpo de mi gato (su forma en este instante) como el límite de los poliedros inscritos en él. Mi gato es lo que es, y sólo porque está ahí, por derecho propio, puedo hablar yo de poliedros inscritos y comprender que éstos pueden aproximarse cada vez más a la forma del gato; y en ese caso, obviamente, su volumen se aproximará al de mi gato. Así, puedo usar la sucesión de esos poliedros abstractos con fines de cálculo; pero como esa sucesión se define como inscrita en mi gato, no podemos usar esas sucesiones abastractas para *definir* ni la forma del gato ni su volumen, *en el viejo sentido de «definir»*.

O, como lo expresa Berkeley, «...por convenientes que parezcan tales analogías o expresiones para facilitar las cuadraturas modernas, descubriremos que no nos iluminan en absoluto respecto a la *naturaleza original real* de las fluxiones; y que a partir de ellas no estamos capacitados para formarnos ideas

justas de las fluxiones consideradas en sí mismas»[48]. Por tanto, los medios que nos facilitan los cálculos no pueden considerarse explicaciones ni definiciones.

Esta era la opinión de todos los matemáticos de la época, incluyendo a Newton y a Leibniz.

Si consideramos ahora el punto de vista moderno, entonces vemos que hemos renunciado por completo a la idea de definición en el sentido en el que la entendían los fundadores del cálculo, así como Berkeley. Hemos renunciado a la idea de una definición que explique el significado (por ejemplo, de una derivada). Este hecho queda velado por nuestra retención del viejo símbolo de «definición» para algunas equivalencias que utilizamos, no para explicar la idea o la esencia de una derivada, sino para *eliminarla*. Y resulta velado por nuestra retención del nombre «cociente diferencial» o «derivada» y el viejo símbolo dy/dx que en un tiempo denotó una idea que ahora hemos descartado. Pues el nombre y el símbolo no tienen ahora otra función que la de servir de etiquetas para el *definiens*, el límite de la sucesión.

Hemos renunciado, pues, a la «explicación» como un caso perdido. La idea intuitiva, descubrimos, llevaba a contradicciones. Pero podemos resolver nuestros problemas sin ella, conservando la mayor parte de la técnica de cálculo, que se basó originalmente en la idea intuitiva. O, más exactamente, *sólo* retenemos esta técnica, en la medida en que es correcta y, con su ayuda, eliminamos la idea. La derivada y la integral se *eliminan*; son sustituidas en efecto por ciertos *métodos* típicos de *calcular* límites.

Pero este cambio ocurrió subrepticiamente y no fue bien entendido por el observador (sobre todo por el observador esencialista) a causa del viejo término «definición» y los viejos nombres y símbolos de la derivada, etc. Esta retención funciona perfectamente: sólo un esencialista puede tener objeciones contra ella.

Es fascinante considerar que todo este admirable desarrollo podría haber sido cortado en flor (como en los días de Arquímedes) si los matemáticos de la época hubieran sido más sensibles a la exigencia de Berkeley —bastante razonable en sí misma— de que nos adhiramos estrictamente a la regla de hablar siempre con sentido.

Sabemos ahora que Berkeley tenía razón cuando, en *The Analyst*, atacó a Newton y a los «matemáticos extranjeros» (como Leibniz) por obtener sus resultados matemáticos en la

[48] Véase Berkeley, *The Analyst*, 1734, sección 47. (Los subrayados son míos.)

teoría de las fluxiones o «en el *calculus differentialis*» por medio de razonamientos ilícitos. Y tenía toda la razón cuando indicó que sus símbolos no tenían sentido. «Nada es más fácil» —escribió— «que idear expresiones y notaciones, para fluxiones e infinitésimos de primer orden y de órdenes posteriores, $\dot{x}, \ddot{x}, \dddot{x}, \ddddot{x}$..., etc. Estas expresiones son, en efecto, claras y distintas y la mente no halla dificultad en concebir que continúan más allá de cualquier límite asignable. Pero si... miramos debajo, si, dejando a un lado las expresiones, nos disponemos a considerar con atención las cosas en sí que se suponen expresadas o señaladas con ello, encontraremos mucha vaciedad, oscuridad y confusión... imposibilidades directas y contradicciones»[49].

Pero los matemáticos de la época no escucharon a Berkeley. Obtuvieron sus resultados y no temieron las contradicciones siempre que sintieron que podían esquivarlas con un poco de habilidad. Esto fue una suerte. Porque el intento de «*analizar el significado*» o de «explicar» sus conceptos no hubiera llevado a ningún sitio, ahora lo sabemos. Berkeley tenía razón; todos esos conceptos carecían de significado en su sentido *y* en el sentido tradicional de la palabra «significado»: eran vacíos, porque no denotaban nada y no simbolizaban nada. Si este hecho se hubiera percibido entonces, el desarrollo del cálculo hubiera vuelto a cortarse de nuevo, como ya se había cortado antes. Fue el descuido de la precisión, el casi instintivo descuido de todo el análisis de singnificados o explicación, lo que hizo posible el espléndido desarrollo del cálculo.

El problema fundamental de toda esta evolución consistió, naturalmente, en conservar el poderoso instrumento del cálculo sin las contradicciones que se habían encontrado en él. No hay duda de que nuestros métodos actuales son *más* exactos que los anteriores. Pero esto no se debe al hecho de que utilicen términos «definidos con exactitud». Tampoco significa que sean exactos: el principal punto de una definición en términos de límites es siempre una aserción *existencial*, y el significado de la frasecita «existe un número» se ha convertido en el centro de perturbación de las matemáticas modernas. Sea lo que fuere lo que uno piense sobre los problemas de la teoría del continuo y del intuicionismo de Brouwer, no puede haber duda de que sólo nos despertamos de nuestros dogmáticos sopores cuando Brouwer nos hizo comprender la diferencia entre pruebas constructivas y pruebas no constructivas; una diferencia cuya importancia admiten hoy incluso sus oponentes[50]. Esto ilustra mi argumento de que el atributo de exactitud no es absoluto y

[49] *The Analyst*, sección 8.

[50] [Véase *Conocimiento objetivo*, *op. cit.*, pp. 125-136. *Ed.*]

de que es inèxacto y muy engañoso usar términos como «exacto» o «preciso» como si tuvieran un significado exacto o preciso.

Mi segundo ejemplo de una definición o explicación es la famosa «teoría de la descripción» de Russell. Su misión aparente era dar una definición o explicación del significado de una frase como «el (tal y tal)», tal como, por ejemplo, «el menor primo mayor que 19». Pero había otros problemas que llevaban a la tarea de definir esta frase. Uno de ellos, como nos dice Russell, era su insatisfacción con las concepciones de Meinong respecto al modo de existencia de las entidades inexistentes[51]. Otro problema —en cierto modo, uno más importante— era mostrar cómo los enunciados que contienen una expresión de la forma «el tal y tal» podían deducirse lógicamente de enunciados que no contenían una expresión de esta forma; por ejemplo, como el enunciado «23 es el menor número primo mayor que 19» podía deducirse de «23 es un número primo y no hay un número primo entre 19 y 23, y $19 < 23$».

La respuesta de Russell al problema se basaba, sin duda, en un análisis del significado intuitivo de los enunciados que contenían la frase «el tal y tal». Pero la definición resultante fue, exactamente igual que en el caso de la definición de la derivada, un medio para *descartar o eliminar el definiendum* intuitivo. Se retuvo el símbolo: pero se convirtió en un símbolo taquigráfico del *definiens*. La cuestión de si el significado querido era adecuadamente representado por el *definiens* pierde utilidad con ello: el símbolo definido no tiene otro significado que ser un símbolo taquigráfico del *definiens*.

Esta es la razón de que la mayor parte de las críticas a la

[51] Véase *The Philosophy of Bertrand Russell*, ed. P. A. Schlipp, 1944, p. 13. Hay que advertir que G. E. Moore, en su ensayo sobre la «Teoría de las Descripciones» de Russell (en el mismo volumen, pp. 177-225) no menciona en ningún sitio ninguno de los problemas que la teoría de Russell se proponía resolver y resolvió, aunque Moore considera uno de sus propios problemas principales el descubrir las razones que pudo haber tenido F. P. Ramsey para descubrir la teoría de Russell como un «*paradigma de la filosofía*». En otras palabras, Moore quiere evaluar la teoría de Russell. Pero nadie puede ni siquiera empezar a comprender el valor de una teoría sin relacionarla con los *problemas* para cuya solución se concibió. Admiro mucho a Moore como realista y como defensor del sentido común. Pero no puede admirarle como analista. Y este ensayo suyo —ahora famoso como *paradigma del análisis*— con sus diversas pequeñas críticas y mejoras de la teoría de Russell, todas ellas sin relación con los *problemas* para cuya solución se concibió esta teoría, ha tenido, mucho me temo, una influencia devastadora. Ha animado a otros a emularla y a ignorar así la máxima de que siempre debemos criticar una teoría en su forma más fuerte posible —en vez de atacar las debilidades de poca importancia en su presentación— si queremos que nuestra crítica sea útil. Con ello ha contribuido a bajar los ya precarios niveles de la discusión filosófica.

definición de Russel (estoy pensando en la contribución de G. E. Moore al volumen de P. A. Schilpp, *The Philosophy of Bertrand Russell*) yerran el tiro, incluso aunque sean quizá más sensibles que Russell al significado intuitivo del término analizado.

Un tercer ejemplo es la definición de *dimensión* (desarrollada por Poincaré, Brouwer, Urysohn y Menger) que tiene cierto papel en *L. I. C.*[52]. El aspecto principal de esta definición es que contribuye a resolver un problema —el problema de la invarianza topológica de la dimensionalidad— que se convirtió en un problema urgente por primera vez por medio de la prueba (de 1878) de Cantor, según la cual los números cardinales de los puntos de un continuo unidimensional y de un continuo de más de una dimensión son los mismos. Naturalmente la definición de la dimensión en términos de teoría de conjuntos llevó a la solución de otros muchos problemas. Considerando el tipo de problema para cuya solución se concibió, el punto más sobresaliente de la definición era que el *definiens* empleaba *términos de la teoría de conjuntos*. Nótese que digo «términos de la teoría de conjuntos» y *no* «términos exactos». De hecho, la exactitud de la teoría de conjuntos intuitiva sobre la que se basaba la definición era algo dudosa en un momento en el que ya existía una teoría de conjuntos más exacta (axiomática). Pero ocurre que el problema particular que esta definición debía resolver no resultaba afectado por las dificultades que llevaron a los desarrollos más exactos; además, la definición puede ser llevada sin cambio a las diversas formas más exactas de la teoría asiomática de conjuntos.

Mi último ejemplo es la definición de Tarski del concepto de verdad. Tarski dice que su única intención era definir la verdad en su sentido absoluto, o, más exactamente, el predicado «... es verdad[ero]» en el sentido de «... corresponde a los hechos» o «... corresponde a la realidad». Con todo, está claro que su intención surgió de la necesidad de resolver algunos problemas muy serios.

Por un lado, la idea de verdad parecía tener una función muy importante. Los filósofos querían decir cosas como «la búsqueda de una teoría verdadera no es lo mismo que la búsqueda de un instrumento útil». Y los lógicos querían decir cosas como, por ejemplo, «si la conclusión de una inferencia válida es falsa, entonces las premisas no pueden ser todas verdaderas». En otras palabras, había una necesidad real de un concepto como el de verdad.

Por otro lado, se afirmaba de modo perfectamente correcto

[52] Véase *L. I. C.*, secciones 32 y 38 ss.

que la frase «es verdad que...» —por ejemplo, «es verdad que 3 + 5 = 8»— era siempre redundante desde un punto de vista lógico y que «3 + 5 = 8» decía exactamente lo mismo; y se señaló (un poco menos correctamente) que «... es verdad» en una frase como «3 + 5 = 8 es verdad» era también una redundancia. (Digo que un poco menos correctamente porque en esta frase se sobreentiende «el enunciado '3 + 5 = 8' es verdadero»; y en él, «es verdadero» no es, obviamente, redundante.) Además de la redundancia sugerida del concepto de verdad, se señaló que el concepto de verdad llevaba a paradojas como la paradoja del mentiroso. Y finalmente, se señaló que la idea de una correspondencia con los hechos o con la realidad era al mismo tiempo inescapable intuitivamente e irremediablemente oscura. Porque, ¿qué queremos decir cuando hablamos de los *hechos* o de la *realidad*? Y ¿qué queremos decir cuando hablamos de una *correspondencia* entre un enunciado o una proposición y un hecho? Wittgenstein había sugerido en el *Tractatus* que esta correspondencia era una similaridad de estructura, que era como la correspondencia entre una melodía y las vibraciones en la aguja del tocadiscos. Pero esta sugerencia equivocada no hizo más que contribuir a la oscuridad.

Esta era, aproximadamente, la situación del problema o, más bien, parte de la situación del problema; y la obra de Tarski consiguió, de la manera más directa, aclarar todos los problemas y oscuridades aquí mencionados. Este es el éxito por el que hay que juzgar su obra. Además, su teoría resultó ser sorprendentemente fuctífera en el campo más técnico de la lógica matemática e incluso en la propia matemática.

Respecto al concepto de verdad intuitivo, normal, cotidiano, la obra de Tarski proporcionó, pienso yo, algo parecido a un análisis de significados valioso. No lo hizo, no obstante, proporcionándonos una definición de este concepto. Por el contrario, Tarski mostró que si introducíamos, sin precauciones especiales artificiales, una definición de verdad en el lenguaje corriente (es decir, en el lenguaje que es «universalista», en el sentido de que en él podemos hablar de cualquier cosa), entonces este lenguaje se hace incoherente. En este sentido, pues, una definición «exacta» de la idea intuitiva de verdad resulta ser imposible. De todos modos, podemos, naturalmente, usar el concepto de verdad con bastante coherencia dentro del lenguaje corriente si aprendemos una lección que puede sacarse fácilmente de la obra de Tarski. La lección, brevemente, es ésta[53].

[53] Véase también mi diálogo, «Autorreferencia y significado en el lenguaje común», *Mind* 63, 1954, pp. 162 ss.; ahora también en *El desarrollo del conocimiento científico (C. & R.)*. [Véase también la discusión de Popper sobre

Debemos distinguir entre *uso* y *mención;* es decir, entre el lenguaje en uso y las partes del lenguaje sobre las que se habla; y debemos recordar que si afirmamos *de* un enunciado que es verdadero, entonces ese enunciado pertenece al lenguaje mencionado, en contraposición al lenguaje en uso —en lenguaje *en* el que afirmamos que un enunciado es verdadero—. Pero esta lección no es lo mismo, desde luego, que «una definición precisa».

Tarski mostró que sólo podía darse una definición precisa con respecto a un lenguaje *formalizado interpretado,* un sistema lingüístico artificial. (El punto sobresaliente es que una definición precisa de una «fórmula bien construida» debe preceder a la definición de un «enunciado verdadero».) Con respecto a los lenguajes formalizados artificiales, Tarski ha dado un *método* de definir la verdad, más que una definición; y es el conocimiento de cómo definir la verdad para esos lenguajes y el conocimiento de que puede darse una definición junto con un conocimiento de las conclusiones que podrían extraerse con su ayuda, lo que hace importante a esta teoría. Este conocimiento, materializado en la teoría de Tarski, es mucho más importante que la definición que da de hecho como ilustración de su método, con respecto a cierto lenguaje muy simple[54].

Podrían ponerse muchos más ejemplos. Pero espero que estos cuatro, junto con el ejemplo de nuestro propio problema de definir el grado de corrobaración, ayudarán a dejar clara mi tesis principal: los problemas de definición o de «análisis» o de «explicación», por sí solos, carecen de significación; sólo pueden tenerla en relación con otros problemas más serios. Nunca puede constituir una empresa racional el sustituir un término por otro más «exacto»; eso no es un problema realmente serio; pero puede convertirse en una tarea seria en relación con un problema serio cuya solución podría facilitarse si tuviéramos una definición o si empezásemos por aclarar ambigüedades terminológicas[55].

Tarski en *El desarrollo del conocimiento científico (C. & R.)* capítulo 10, y *Conocimiento objetivo (Obj. Kn.)*, capítulos 2 y 9. *Ed.*]

[54] La obra de Tarski ha fomentado la concepción de que tenemos que construir un lenguaje modelo artificial siempre que queramos dar una definición «exacta». No creo que esta concepción interprete como es debido la obra de Tarski que se refiere a *todos* los lenguajes de un conjunto muy extenso.

[55] Una definición útil —que nos ayude a resolver problemas reales— será siempre una definición *eliminatria* más que *explicativa*, como he indicado antes; aunque hay que destacar que nuestra investigación se guiará aquí, como siempre, por la intuición. Pero la intuición de significados que puede guiarnos a proponer una cierta definición es un asunto privado nuestro; y la aceptación de la propuesta dependerá de su fertilidad, de la ayuda que nos preste en la solución de nuestros problemas. Esta distinción entre definiciones *a*) elimina-

Estoy dispuesto a admitir que un aumento de precisión puede tener gran importancia. No por sí mismo, sino porque puede aumentar la *constrastabilidad* de una teoría. (*Cf.* sección 37 de *L. I. C.*) En otros casos puede ayudarnos a hacer precisamente las distinciones necesarias para resolver un problema serio. Pero la precisión no debe convertirse en un fetiche. Tampoco debe tratarse como un valor en sí misma, como puedan serlo la claridad o la lucidez.

A propósito, hay toda una categoría de problemas cuya solución parece que podría ser impulsada por una definición, mientras que las definiciones, en realizada, no harían más que privarlos de su carácter empírico al cambiar un problema factual en un problema verbal. Por ejemplo, si queremos resolver el problema reconocidamente confuso, pero tal vez factual «¿son todas las obras de arte bellas o hay obras de arte que no son bellas?» tendríamos que cuidarnos mucho de la tentación de definir «obra de arte» o «bella» y tendríamos que tratar, en cambio, de recordar cosas que podrían describirse razonablemente como obras de arte y al mismo tiempo como no bellas. En otras palabras, deberíamos buscar ejemplos; *Pierrot Lunaire* de Schönberg; *1984* de Orwell; algunas caricaturas (incluso de Leonardo); y Gulliver en Brobdingnag, quizá puedan ser candidatas. Esta lista es obviamente tan esquemática como subjetiva, y la solución basada en ella es, por tanto, dudosa. Pero podría mejorarse. Además, lo vago y lo dudoso de la solución encajan perfectamente con lo vago y lo dudoso del problema. Ahora bien, si, por otro lado, definimos, por ejemplo, «obra de arte» de modo que incluya entre sus propiedades la belleza, podemos estar seguros de resolver nuestro problema con «precisión». Pero al hacerlo sustituimos (quizá sin advertirlo) nuestro problema factual por *un problema puramente verbal* y ninguna reflexión sobre obras de arte reales puede contribuir ya a su solución. Mi lista no será ya esquemática ni subjetiva, pero tampoco será pertinente en absoluto; podemos decir *a priori* que cada uno de sus elementos o tiene que ser bello, contrariamente a mi sugerencia, o si no, no puede ser una obra de arte, contrariamente también a mi sugerencia. Y podemos afirmar esto incluso aunque no conozcamos ninguna de las obras mencionadas en esta lista u otra similar. (En el terreno científico, se logra una transformación muy similar de los problemas factuales en problemas puramente verbales por medio del con-

---

torias y *b*) explicativas, se corresponde estrechamente con lo que he descrito en otro lugar como *a*) definiciones de derecha a izquierda y *b*) definiciones de izquierda a derecha, o esencialistas o aristotélicas [véase *La sociedad abierta (O. S.)*, capítulo II, sección II]. [*Cf. Búsqueda sin término (U. Q.)*, p. 34.]

vencionalismo y por medio de todas las interpretaciones o estratagemas convencionalistas; *cf.* secciones 19 y 20 de *L. I. C.* En filosofía, la transformación de problemas filosóficos o de otra clase en problemas verbales es simplemente una maldición, sobre todo desde el día en que definición y «explicación» se convirtieron en el objetivo declarado de la filosofía «analítica».)

Mis ejemplos pueden ayudar a poner de relieve una lección que nos enseña toda la historia de la ciencia; que la exactitud absoluta no existe, ni siquiera en lógica y en matemátcas (como la ilustra el ejemplo de la historia, inacabada aún, del cálculo); que nunca debemos intentar ser más exactos de lo necesario para la solución del problema que nos ocupe; y que la exigencia de «algo más exacto» no puede constituir, en sí misma, un problema (excepto, naturalmente, cuando la mayor precisión mejore la contrastabilidad de alguna teoría).

En lo que concierne al análisis de significados por sí mismo, es siempre *inútil*. Donde dos o tres sentidos de una palabra están habitualmente confundidos (como, por ejemplo, en el caso de «probabilidad») es aconsejable llamar la atención sobre el hecho mostrando que hay enunciados que son válidos para uno u otro sentido y que se hacen incompatibles si los sentidos no se *distinguen*. Pero, de otro modo, preguntas como «¿qué es corroboración?» o «¿qué significa 'corroboración'?» son vacías[56].

---

[56] Creo que esta observación aclara la llamada «paradoja del análisis»; porque a la luz de mi observación, puede interpretarse que esta «paradoja» muestra que un análisis (o una definición que no introduzca una breve etiqueta para una larga historia) es *siempre inútil*, excepto si el tipo de término que *va a admitirse* en el «*analysans*» (o *definiens*), está indicado por un *problema determinado*: el problema a cuya solución se supone que contribuye el análisis.

La paradoja del análisis normalmente ilustrada por la definición analítica «un hermano es un hijo varón de los mismos padres» [«a brother is a male sibling»], es un ejemplo bastante bueno y uno que cuenta con la aprobación del propio Moore. He encontrado algunos ejemplos que son quizá más instructivos y más divertidos, aunque tienen una intención sorprendentemente similar, en un folleto de amplia circulación editado por la Auto-escuela de la Policía Metropolitana y publicado por Her Majesty's Stationery Office en 1955. El título es *Roadcraft: A Manual of Driving Instruction for Students of the Motor Car Wing* (Destreza en la carretera: Un manual de aprender a conducir para los alumnos de la sección de automóviles) y se ocupa mucho de la explicación y el análisis de significados, como lo muestra esta cita, bastante representativa:

> 5. La concentración puede definirse como la plena aplicación de la mente y el cuerpo a una tarea determinada, con la total exclusión de todo lo que no se refiera a dicha tarea (p. 7).

A los filósofos de la moral puede interesarles el análisis de «buen criterio» que contiene el siguiente párrafo:

> 8. Con ayuda de la VISTA, OIDO, BUENAS CONDICIONES GENERALES Y CONCENTRACION, un conductor será capaz de

ejercer su buen CRITERIO, que es la capacidad de distinguir entre lo correcto y lo equivocado, lo bueno y lo malo, lo prudente y lo imprudente (pp. 7 s.)

Las dos citas siguientes son ejemplos de definiciones directas:

15. Esfuerzo en este sentido, puede definirse como el empleo de fuerza (p. 11).

18. Patinar puede definirse como sigue: *Movimiento involuntario del coche debido a que el agarre de los neumáticos a la carretera se hace menor que la fuerza* o fuerzas que actúan sobre el coche (p. 18).

Creo que ningún estudioso serio de la historia de las ideas pondrá en duda mi afirmación de que la filosofía y los filósofos son los responsables últimos de esta clave de verborrea vacía, y no el Comisario de Policía.

PARTE II

LA INTERPRETACION PROPENSIVISTA DE LA PROBABILIDAD

CAPITULO PRIMERO

PROBABILIDADES OBJETIVAS Y SUBJETIVAS

En este capítulo del *Post Scriptum* me propongo desarrollar con bastante extensión la distinción (hecha en la sección 48 de *L. I. C.*) entre las interpretaciones objetivas y subjetivas del cálculo de probabilidades [1].

La teoría subjetiva de la probabilidad surge de la creencia de que utilizamos la probabilidad sólo si tenemos un *conocimiento insuficiente.* Critiqué esta teoría en muchos lugares, en *La lógica de la investigación científica:* pero, como era de esperar, sobrevivió a mis críticas. Incluso entre los físicos, la teoría sigue siendo muy popular. A pesar de ello, en mi opinión, esta teoría ha llevado a mucha confusión y a mucho razonamiento inválido dentro de la física. Por esta razón voy a renovar mi ataque con el propósito de mostrar que debe ser abandonada a pesar de su plausibilidad *prima facie.*

El ataque a la interpretación subjetiva es el tema principal de este capítulo y de los que le siguen. Pero estos capítulos van precedidos y seguidos de una alegación en favor de una nueva

---

[1] [Con el material que se presenta en este capítulo y el siguiente, compárense las siguientes publicaciones de K. R. Popper: «Two Autonomous Axiom Systems for the Calculus of Probabilities», *The British Journal for the Pilosophy of Science* 6, 1955, núm. 21, pp. 51-57 (véase también 6, núm. 22, p. 176, y núm. 24, p. 351, donde se hacen notar ciertas erratas); «The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory» en Stephan Körner ed., en colaboración con M. H. L. Pryce, *Observation and Interpretation. Proceedings of the Ninth Symposium of the Colson Research Society, University of Bristol.*) Londres, 1957, pp. 65-70 y 88-69 (para las erratas, véase *The British Journal for the Pilosophy of Science* 8, 1958, núm. 32, p. 301, n. 1); «Probability Magic, or Knowledge out of Ignorance», *Dialectica* 11, 1957, núm. 3/4, pp. 354-372; «On Mr. Roy Harrod's New Argument for Induction», *The British Journal for the Pilosophy of Science* 9, 1958, núm. 35, pp. 221-224; «The Propensity Interpretation of Probability» *The British Journal for the Pilosophy of Science* 10, 1959, núm. 37, pp. 25-42; «Probabilistic Independence and Corroboration By Empirical Tests», *The British Journal for the Pilosophy of Science* 10, 1960, núm. 40, pp. 315-318; «Creative and Non-Cretive Definitions in The Calculus of Probability» *Synthese* 15, núm. 2, junio 1963, pp. 167-186 (para las erratas, véase *Synthese* 21, núm. 1, 1970, p. 107).

Véase también Popper, «On Rules of Detachment and So-called Inductive Logic» y «Theories, Experience and Probabilistic Intuitions», en I. Lakatos, ed., *The Problem of Inductive Logic,* 1968, pp. 130-144, pp. 285-303. *Ed.*]

interpretación objetiva de la probalidad a la que llamo «interpretación propensivista». Esta interpretación marca, quizá, el cambio más significativo en mis concepciones desde 1934 y la expondré con gran detalle también en los tomos II y III del *Post Scriptum* [2].

1. LOS SENTIDOS DE LA PROBABILIDAD [3]

Es un error creer que la palabra «probabilidad» sólo tiene un sentido importante o, como mucho, dos. La palabra tiene muchos sentidos, algunos de los cuales tienen poco en común. Propongo que usemos la palabra «probabilidad» aquí y en otros lugares, para todos los sentidos que satisfagan el bien conocido *cálculo matemático de probabilidades* y sólo para ellos y que introduzcamos nuevos términos —tales como «aceptabilidad», «grado de corroboración» o «grado de confirmación», etc.— en todos aquellos casos en los que la palabra «probabilidad» pueda emplearse con perfecta propiedad en el uso filosófico u ordinario pero en los que *se ha puesto en duda* (por ejemplo, por mí mismo en *L. I. C.*) *que este uso esté de acuerdo, de hecho, con las leyes del cálculo de probabilidades.* (Para las leyes de dicho cálculo, véanse principalmente los Apéndices *III a *V de mi *L. I. C.*)

---

[2] [Cuando Popper presentó por primera vez esta teoría propensivista en un artículo que leyó, en su ausencia, su discípulo Paul K. Feyerabend (véase S. Körner, ed. *Observation and Interpretation, op. cit.*) R. B. Braithwaite comparó la idea de Popper de la propensión, cosa interesante, con el «hábito» o el «aspirar a» («would-be») de C. S. Peirce; y desde entonces, la teoría de Popper se ha atribuido a Peirce. Pero hay, naturalmente, importantes diferencias. Por ejemplo, del mismo modo que un campo de fuerzas puede estar físicamente presente incluso cuando no hay un cuerpo (de contrastación) sobre el que pueda actual, igual puede existir una propensión a que una moneda caiga mostrando la cara aunque caiga una sola vez y en esa ocasión muestre la cruz. De hecho, puede haber una propensión sin ninguna caída de la moneda. La diferencia más importante con la teoría de Peirce, resaltada por Popper en 1957, es su teoría *relacional* de la propensión. Puede mencionarse que la única crítica avanzada por D. H. Hellor (*A Matter of Chance*, 1971), que adopta la terminología de Popper, es que no está de acuerdo con la teoría relacional de Popper (aparentemente por razones de hábitos del lenguaje) y que vuelve, pues, a la concepción de Peirce. Pero la concepción relacional constituye, evidentemente, un perfeccionamiento considerable: el riesgo de muerte de una persona enferma puede reducirse mucho (o incrementarse mucho) con el invento de una nueva terapia. Y la propensión (vida media) de un núcleo atómico a desintegrarse puede aumentar radicalmente, por ejemplo, bombardeándolo con neutrones lentos. *Ed.*]

[3] Las ideas fundamentales de esta sección y la siguiente se remontan a *L. I. C.*, sección 48. La mayor parte de las innovaciones se introdujeron en mi nota en *Mind*, 1938, reimpresa en el Apéndice *II de *L. I. C.*

El grado de corroboración no volverá a tratarse aquí (véase Parte I, capítulo IV). Sólo trataré de los conceptos de probabilidad que satisfagan incuestionablemente el cálculo formal de probabilidades. (Pero no se tratará de todos esos conceptos.)

Si se acepta mi propuesta terminológica, entonces una frecuencia relativa, por ejemplo, se llama correctamente «probabilidad»; porque sobre la base de la definición dada en *L. I. C.* (Sección 52 y Apéndices *II y *VI), las frecuencias relativas satisfacen tautológicamente todas las leyes del cálculo de probabilidades.

Pero la frecuencia relativa no es, en modo alguno, la única interpretación que satisface este cálculo. Hay muchas interpretaciones más. De ellas, las siguientes tienen la mayor importancia: la interpretación según la teoría de conjuntos; la interpretación de la probabilidad como medida (ponderada) de las posibilidades y, más especialmente, la interpretación propensivista que considera que esas posibilidades son propensiones físicas (que pueden considerarse tan reales físicamente como las fuerzas físicas); y la interpretación lógica.

Ahora bien, en lo que he cambiado de opinión desde 1934 —es decir, desde la publicación de *L. I. C.*— es principalmente, en el punto siguiente: Mientras que anteriormente yo creía que la interpretación frecuencial era fundamental para la comprensión de las *teorías físicas* probabilísticas (y de las teorías de juegos de azar), ahora creo que es más adecuada la interpretación propensivista.

Sin embargo, antes de explicar este cambio de opinión, voy a tratar sobre dos fórmulas muy sencillas para dejar totalmente claro qué tipo de formalismo es el que tenemos que interpretar.

2. PROBABILIDADES ABSOLUTAS Y RELATIVAS

Sólo hay dos fórmulas a las que hay que referirse en una discusión de las interpretaciones del cálculo de probabilidades; la primera es

R) $$p(a,b) = r$$

o, expresado con palabras, «la probabilidad (relativa) de a dado b es igual a r», donde r es una fracción entre 0 y 1 (ambos límites inclusive).

La segunda fórmula es

A) $$p(a) = r$$

o, expresado con palabras, «la probabilidad (absoluta) de a es igual a r».

La probabilidad relativa de *a* dado *b* se ha llamado también algunas veces «la probabilidad condicional de *a* bajo la condición *b*»; y la probabilidad absoluta de *a* se ha llamado a veces probabilidad «previa» o «inicial» o *a priori* de *a*.

Si interpretamos R) en el sentido de la *interpretación frecuencial* que se expuso con detalle en *L. I. C.*, entonces

$$p(a,b) = r$$

se convierte en «la frecuencia relativa de *a* dentro de la clase de referencia *b* (o la sucesión de referencia *b*) es igual a *r*».

Por otro lado,

$$p(a) = r$$

difícilmente puede interpretarse en términos frecuenciales (excepto de un modo trivial, dando por «sobreentendida» la clase de referencia *b*). Porque no tiene sentido decir «los gorriones son frencuentes» a menos que se sobreentienda «entre los pájaros europeos» o algo parecido. Pero

$$p(a) = r$$

puede interpretarse sin dificultad en términos diferentes de las frecuencias, por ejemplo, dentro de lo que llamo la «interpretación lógica» de la probabilidad. Aquí consideramos que las letras «*a*», «*b*», etc., son nombres de enunciados. La probabilidad lógica absoluta de *a* —es decir, *p(a)*— es entonces lo que describí en *L. I. C.* como la «probabilidad lógica»: su valor *r* es mayor cuanto menos diga el enunciado *a*. O, dicho de otro modo, cuanto mayor sea el contenido de *a*, menor será el valor de su probabilidad lógica absoluta. Esto lo justifica el hecho de que la aserción *a* (mañana lloverá) es claramente más probable que la aserción *ab* (lloverá mañana y saldrá el sol el sábado) siempre que *b* no se deduzca de *a* (lo que ocurre, por ejemplo, si es una tautología).

Se ha dicho en ocasiones que la idea de probabilidad absoluta carece de sentido; y, en efecto, es difícil darle un significado útil dentro de la interpretación frecuencial, como hemos visto. No obstante, *p(a)* difícilmente puede describirse como carente de sentido, incluso dentro de la teoría frecuencial, pues el sentido de la probabilidad absoluta puede definirse con ayuda de la probabilidad relativa.

Para mostrarlo, podemos recordar que todas las teorías de la probabilidad operan con la idea de la conjunción o el producto o la intersección de *a* y *b*, representado aquí por «*ab*» (léase «*a* y *b*») y con la idea de la negación o el complemento de *a*, representado aquí por $\bar{a}$ (léase «no-*a*»). Si se introducen esas dos ideas, entonces $a\bar{a}$ se convierte en un elemento del sis-

tema (suceso, clase o enunciado, etc.) auto-contradictorio o vacío; y la negación de $a\bar{a}$, es decir $\overline{a\bar{a}}$, se convierte en un elemento «lleno» o tautológico, por ejemplo, un enunciado tautológico. Podemos convenir en escribir «t» (es decir, tautología) en vez de $\overline{a\bar{a}}$. Entonces podemos definir la probabilidad absoluta como sigue:

DA) $$p\,(a) = p(a,t).$$

Como hemos definido ahora la probabilidad absoluta en términos de la probabilidad relativa, tiene que «tener sentido». Y si suponemos un universo finito de discurso —digamos que con n elementos— el significado de

$$p\,(a) = r \qquad (0 \leq r \leq 1)$$

es perfectamente claro desde el punto de vista de la frecuencia relativa: $p\,(a) = r$ afirma que entre los n elementos del universo, nr elementos tienen la propiedad a. La creencia de que $p\,(a)$ carece de sentido en la teoría frecuencial se debe al hecho de que, por regla general, no suponemos que el universo en el que realmente vivimos sea finito. (Al menos, por lo general, no suponemos que sea finito en el tiempo.) Y en un universo infinito, el sentido de «$p\,(a) = r$» es, en efecto, muy oscuro, porque si no es ni 1 ni 0, en general será indeterminado en el sentido matemático.

En un universo finito también podemos definir la probabilidad relativa en términos de la probabilidad absoluta por medio de la definición

DR) $$p\,(a,b) = p(ab)/p(b).$$

Esta definición puede extenderse a un universo infinito, pero sólo con la condición de que $p\,(b) \neq 0$. [En un universo finito podemos decir $p\,(a,b) = 1$ siempre que $p\,(b) = 0$.]

A continuación voy a operar principalmente con probabilidades relativas; es decir, voy a investigar diversas interpretaciones de

$$p\,(a,b) = r$$

Si se toman las probabilidades absolutas podemos considerarlas como definidas por la probabilidad relativa al modo de DA).

No supongo aquí que se tenga ningún conocimiento del verdadero cálculo. En *L. I. C.*, Apéndices *IV y *V, se encuentran una exposición de sus axiomas y de las derivaciones elementales.

3. LA INTERPRETACION PROPENSIVISTA. INTERPRETACIONES OBJETIVAS Y SUBJETIVAS

En esta sección me propongo introducir lo que he llamado la interpretación propensivista, pero sin discutirla con gran detalle y sin exponer mis principales argumentos en su favor. (Estos se expondrán en la sección 20.)

La interpretación propensivista puede considerarse en relación con la interpretación clásica, que define la probabilidad como el número de casos favorables dividido entre el número de casos posibles. Esto sugiere que podemos interpretar la probabilidad como una *medida de las posibilidades*. Si, y sólo si, puede contarse el número de posibilidades y si todas esas posibilidades son iguales, tenemos un caso al que es claramente aplicable la definición clásica. Así, la interpretación de la probabilidad como una medida de las posibilidades es una generalización natural de la definición clásica.

Ahora bien, la interpretación propensivista está estrechamente relacionada con la interpretación que considera la probabilidad como una medida de las posibilidades. Lo único que añade a esto es una interpretación física de las posibilidades, a las que no considera como meras abstracciones, sino como tendencias o propensiones físicas a producir el estado posible de cosas, tendencias o propensiones a realizar lo que es posible. Y supone que la fuerza relativa de una tendencia o propensión de este tipo se expresa en la frecuencia relativa con la que logra realizar la posibilidad en cuestión.

Así pues, las frecuecias relativas pueden considerarse como los resultados o las expresiones exteriorizadas, o las apariencias, de una disposición o tendencia o propensión física oculta y no observable directamente; y una hipótesis relativa a la fuerza de esta disposición o tendencia o propensión física puede contrastarse por medio de contrastaciones estadísticas, esto es, por medio de observaciones de frecuencias relativas.

Esta es, brevemente descrita, la interpretación propensivista. Se me preguntará por qué propongo introducir propensiones ocultas tras las frecuencias. Mi respuesta es que conjeturo que esas propensiones son físicamente reales en el mismo sentido que las fuerzas de atracción o repulsión, por ejemplo, pueden ser físicamente reales. Aparte de esto, trataré de mostrar que puede resultar muy provechoso el operar en física con estas propensiones (e incluso con campos de propensiones que son análogos a los campos de fuerzas). Estimo ahora que la interpretación frecuencial es un intento de prescindir de la realidad física oculta, un intento que mereció la pena realizar y que

puede aplicarse con éxito, pero que no logra un éxito completo en todos los campos. Creo que tendrá que dar paso, al final, a la interpretación propensivista.

Considerada desde un punto de vista formal, la diferencia entre ambas interpretaciones es ésta: la interpretación frecuencial puede explicar *enunciados de probabilidades singulares* —tales como «existe una probabilidad de 1/2 de que en la próxima tirada con esta moneda salgan caras»— como singulares sólo gramaticalmente (o «formalmente»), como traté de explicar en *L. I. C.*, sección 71. La interpretación frecuencial atribuye al suceso único una probabilidad *solamente* en la medida en que este suceso único es un elemento de una sucesión de sucesos con una frecuencia relativa.

Contrariamente a esto, la interpretación propensivista atribuye una probabilidad a un suceso único como representante de una sucesión *virtual o concebible* de sucesos, más que como un elemento de una sucesión real. Atribuye al suceso *a* una probabilidad *p (a,b)* considerando *las condiciones que definirían a esta sucesión virtual:* éstas son las condiciones *b*, las condiciones que producen la propensión oculta y que dan al caso único una cierta probabilidad numérica. Sólo en caso de que queramos *contrastar* la probabilidad númerica atribuida, tendremos que realizar un segmento de la sucesión virtual lo suficientemente largo como para que nos sea posible aplicarle una contrastación estadística significativa.

Así pues, la principal diferencia entre la interpretación frecuencial y la propensivista estriba en la posición de los enunciados de probabilidades singulares. En la teoría frecuencial tienen un papel periférico, pero en la interpretación propensivista tienen un papel central, ya que esta interpretación ve cada caso como el resultado de una propensión, por así decirlo, o quizá como el resultado de propensiones encontradas, aunque éstas sólo puedan contrastarse estadísticamente.

Más adelante tendré mucho que decir sobre la interpretación propensivista. En el contexto actual sólo quería presentar la idea. Sólo me queda añadir (en la próxima sección) una breve observación sobre el modo en que esta interpretación encaja en mi programa general de proponer hipótesis y de contrastarlas mediante experimentos.

Con el fin de tener un nombre común para las interpretaciones frecuencial y propensivista (y tal vez para otras interpretaciones similares que puedan proponerse todavía), las llamaré interpretaciones «*objetivas*», extendiendo así en cierto modo, el ámbito de una palabra que introduje en *L. I. C.*, sección 48. Las interpretaciones objetivas suponen que la probabilidad de que salga cara en una tirada depende solamente de las condiciones

físicas u otras similares y *no* del estado de nuestro conocimiento. Y llamaré «*sujetivas*» a aquellas interpretaciones que consideran que la probabilidad de que salga cara depende del estado de nuestro conocimiento (subjetivo) o quizá del estado de nuestras creencias.

4. CONTRASTACIONES EXPERIMENTALES Y SU REPETICION: INDEPENDENCIA

Toda teoría científica supone que en ciertas condiciones ocurrirán ciertas cosas. Toda contrastación consiste en un intento de crear esas condiciones y averiguar si podemos obtener un contra-ejemplo, incluso aunque esas condiciones se hayan creado; por ejemplo, variando otras condiciones que no se mencionen en la teoría. (Esto muestra, por cierto, que no debe añadirse a una teoría una cláusula *ceteris paribus* como «siendo iguales todas las cosas», puesto que destruiría la contrastabilidad de la teoría.)

Este procedimiento fundamentalmente claro y sencillo para la contrastación experimental puede aplicarse, en principio, a las hipótesis probabilísticas del mismo modo en que puede aplicarse a hipótesis *no probabilísticas* o, como podemos decir, en pro de la brevedad, hipótesis «*causales*».

Resalté este punto incluso cuando defendía la interpretación frecuencial de las hipótesis probabilísticas (físicas). Expliqué entonces que hipótesis probabilísticas hacen afirmaciones sobre la distribución estadística de ciertas propiedades de los sucesos dentro de *clases o sucesiones* de sucesos y que esas afirmaciones hipotéticas pueden contrastarse de un modo análogo a la contrastación de las hipótesis causales.

Ahora me siento inclinado a decir que la relación entre hipótesis probabilistas y causales y la manera de contrastarlas es incluso más similar de lo que yo pensaba anteriormente: desde el punto de vista de la interpretación propensivista, las hipótesis causales pueden interpretarse fácilmente como hipótesis que afirma una propensión igual a 1 («certeza»).

Esta interpretación concuerda con el carácter de las hipótesis causales, que dan lugar a predicciones *singulares*; porque la interpretación propensivista atribuye probabilidades también a los «sucesos singulares», es decir, a las «ocurrencias» (en el sentido de *L. I. C.*, sección 23).

Por tanto, contrastamos las hipótesis, sean probabilísticas o causales, creando ciertas condiciones y viendo si producen el efecto (singular) predicho si se varían otras condiciones. La hipótesis probabilística predice que el suceso singular tiene una

cierta propensión a realizarse. Esta predicción puede contrastarse *repitiendo* el experimento bajo las condiciones prescritas y anotando la distribución de la frecuencia en repetidos experimentos. Pero también la hipótesis causal se contrasta por medio de experimentos repetidos: y aquí la corroboración exige que las frecuencias relativas de los resultados dentro de la sucesión de contrastaciones repetidas iguales a la unidad (o, en la práctica, muy próximas a ella).

La contrastación por medio de experimentos tiene, pues, dos aspectos: la variación de las condiciones es uno; y mantener constantes las condiciones que se mencionan como pertinentes en la hipótesis es el otro, el único aspecto que nos interesa aquí. Es decisivo para la idea de *repetir un experimento.*

El experimento a repetir se define, o describe, por esas condiciones. Es, por tanto, absolutamente esencial para la repetición de un experimento que cada repetición tenga lugar bajo las mismas condiciones enunciadas. Pero esto significa que en una disposición experimental, *los primeros experimentos no deben afectar a los posteriores.* Porque, de otro modo, los posteriores operan bajo nuevas condiciones. En otras palabras, los experimentos tienen que ser *independientes.* Esto forma parte de la idea de repetición y no tiene nada en absoluto que ver con la cuestión de si la hipótesis que está contrastándose es probabilística o causal.

Para expresar este punto en términos prácticos: sea que nuestro experimento termine con la lectura de un medidor consistente en una aguja que recorre una escala. Si el aparato no está bien engrasado, puede quedarse atascado y, por esa razón, una repetición del experimento puede terminar con la misma lectura que el experimento anterior. Está claro que estamos ante un «efecto secundario» o *dependencia* (no-independencia) de los experimentos.

En consecuencia, si nos interesan los experimentos que incluyen efectos secundarios (tales como el movimiento browniano o el aprendizaje de memoria), entonces consideraremos, desde el punto de vista de la contrastación de nuestra teoría, los experimentos que presentan efectos secundarios junto con el experimento original como parte de un experimento *singular*, en vez de como repeticiones del mismo experimento; y sólo consideraremos repetición auténtica un experimento que comience por el principio y que incluya, a su vez, el estudio de los efectos secundarios.

En resumen, es esencial para una sucesión de experimentos repetidos, que cada uno se lleve a cabo bajo las mismas condiciones enunciadas y que, por esa misma razón, los experimentos sean *independientes*, lo que no es más que otro modo de

decir que los experimentos deben estar libres de efectos secundarios.

Las contrastaciones de las hipótesis probabilísticas más simples implican esas sucesiones de experimentos repetidos y, por tanto, independientes de la misma manera que las contrastaciones de hipótesis causales. Y la probabilidad o propensión hipotéticamente estimada se contrastará por medio de las distribuciones de frecuencias en esas sucesiones independientes. (La distribución de las frecuencias de una sucesión independiente debe ser «normal» o «gaussiana»; y como consecuencia, tiene que indicar claramente si la propensión conjeturada debe considerarse o no como refutada o corroborada por la contrastación estadística.)

Ahora bien, la independencia, en el sentido mencionado, ha sido durante mucho tiempo uno de los conceptos fundamentales de la teoría de la probabilidad. En esta teoría se explica como sigue: los últimos experimentos son independientes de los primeros si, y sólo si, la probabilidad de obtener un cierto resultado no está influenciada por los experimentos anteriores; o dicho de otro modo, si las probabilidades continúan siendo las mismas para cada uno de los experimentos a lo largo de toda la sucesión.

Esto denota claramente la idea de que las probabilidades dependen de las condiciones experimentales, de la disposición experimental. Lo que se supone, en esta teoría de la independencia, es que si la disposición experimental de los experimentos posteriores es la misma que la de los anteriores, entonces, las probabilidades de obtener resultados serán también las mismas; y viceversa, si las probabilidades permanecen iguales después de cada experimento, entonces los experimentos no tienen efectos secundarios: no cambian las condiciones.

Para ser más exacto tendría que haber dicho «no cambian las condiciones *pertinentes*». Porque evidentemente, el hecho de que se haya efectuado otro experimento antes y haya producido un cierto resultado, es parte de «las condiciones» del mismo modo que el hecho de que se haya desatado una tempestad de lluvia en Borneo o una tempestad de indignación en Alejandría mientras se lanzaba una moneda al aire en Londres. La cuestión es que esas condiciones no tienen ningún efecto sobre la probabilidad de que salga cara: son *no pertinentes*: no influencian el conjunto de condiciones *especificadas y pertinentes* y, por tanto, no influencia la probabilidad.

Así, se dice que un experimento es «*independiente*» de otro, o de ciertas condiciones, o que no resulta afectado por esas condiciones si, y sólo si, éstas no cambian la probabilidad del resultado. Y las condiciones que no tienen ningún efecto sobre

la probabilidad del resultado reciben el nombre de «*no pertinentes*».

Las definiciones de estos términos con ayuda del cálculo formal son muy sencillas. Sea *a* el resultado en cuya probabilidad estamos interesados y *b* las condiciones experimentales pertinentes. Entonces

$$p\ (a,b) = r$$

será la aserción según la cual bajo las condiciones *b*, el resultado *a* tiene una probabilidad *r*. (En la interpretación propensivista, es la aserción según la cual las condiciones *b* producen una propensión *r* a que se realice el resultado *a*.)

Ahora, sea *c* alguna condición inicial, por ejemplo, la condición de que el experimento en cuestión se realice después de otros dos experimentos (que habrán dado, ambos, el resultado *a*, por ejemplo). Entonces diremos que esas condiciones son no pertinentes, o que el resultado *a* es independiente de ellas, si la probabilidad de *a* continúa siendo, bajo las condiciones *bc* (esto es, la conjunción de ambas condiciones), la misma que era bajo la condición *b* sola. Es decir, si, y sólo si

$$(*) \qquad p\ (a,bc) = p(a,b)$$

Para ser más exactos tenemos que mencionar *b* en nuestras formulaciones verbales también: en este caso decimos que *a* es independiente de *c* dado *b*; y que *c* no es pertinente a *a*, dado *b*.

La independencia de los experimentos o sucesos y la no pertinencia de las condiciones son, como se verá ahora, un conjunto correlativo de términos que expresan *una* idea. El siguiente cálculo de probabilidades pone de relieve este hecho: si $p\ (a,b) > 0$ y $p\ (c,b) > 0$, entonces $p\ (a,bc) = p(a,b)$ si, y sólo si, $p\ (c,ba) = p(c,b)$; o, expresado verbalmente, *a* es independiente de *c* dado *b* (y *c* no es pertinente a *a* dado *b*), si, y sólo si, *c* es a su vez independiente de *a* dado *b* (y *a* es no pertinente a *c*, dado *b*). Por esta razón no hay necesidad de distinguir entre independencia y no pertinencia de las condiciones.

5. LA INTERPRETACION LOGICA

Lo mejor es considerar una fórmula como

$$p\ (a,b) = r$$

como un enunciado o una aserción sobre los objetos *a* y *b*. Los símbolos «*a*» y «*b*» son así *nombres* (variables) de los objetos sobre los que estamos hablando. Estos objetos pueden ser, por

ejemplo, *sucesos* o clases o conjuntos de sucesos; o pueden ser, por ejemplo, *enunciados* (o quizá conjuntos de enunciados, tales como teorías deductivas).

Hay un tipo de transición bastante directa de una interpretación que considera a *a* y *b* como enunciados: sencillamente, podemos escoger los enunciados que describen los sucesos. De ese modo procedemos de una interpretación a la otra de un modo que no afecta a lo que en realidad queremos afirmar cuando afirmamos «*p (a,b) = r*». Sólo nuestro método para afirmarlo, nuestro método de expresión, habrá cambiado. Y si el enunciado «*p (a,b) = r*» era un enunciado factual para empezar, por ejemplo una estimación de una probabilidad hipotética, en vez de una lógica o tautológica, entonces, naturalmente, conservará ese carácter después de su reinterpretación.

Pero si consideramos nuestros argumentos «*a*», «*b*», «*c*»... como nombres de enunciados, entonces *p (a,b) = r* puede interpretarse también de un modo muy diferente. En lo que he llamado (en *L. I. C.*, sección 48) *la interpretación lógica de la probabilidad,* «*a*» y «*b*» se interpretan como nombres de *enunciados* (o proposiciones) y

$$p\,(a,b) = r$$

como una afirmación sobre el *contenido* de *a* y *b* y sobre su *grado de proximidad lógica*; o, más exactamente, sobre el grado en el que el enunciado *a* contiene información que está contenida en *b*. Si *b* dice todo lo que dice *a*, de modo que *a* se sigue de *b*, entonces *p (a,b)* = 1. Si *b* es coherente y es contradicho por *a*, entonces, *p (a,b)* = 0, mientras que, naturalmente, *p (a,ab)* = 1. Si *a* no se sigue de *b* ni contradice a *b*, entonces el valor de *p (a,b)* estará en algún punto entre 0 y 1 (incluidos ambos extremos); y será próximo a 1 si *a* dice sólo un poco más de lo que dice *b*, y próximo a 0 si dice cosas muy diferentes de las que dice *b*.

Esta «interpretación lógica» considera el cálculo de probabilidades como una generalización de la lógica ordinaria, por así decirlo. Puede justificarse o bien con argumentos intuitivos o, de un modo más completo, con argumento formales.

La justificación intuitiva es como sigue. Sea *a* el enunciado «Sócrates es mortal» y *b* el enunciado «todos los hombres (o el 100 por cien de los hombres) son mortales y Sócrates es un hombre»; entonces diremos que *p (a,b)* = 1, porque *a* se sigue de *b* y, en efecto, dado *b*, podemos considerar a *a* como cierto.

Pero sea *a* el mismo enunciado que antes y *b* el enunciado «el 92 por ciento de todos los hombres son mortales y Sócrates es un hombre», entonces, *a* no será *cierto*, dada la información *b* pero será sumamente *probable*; y podemos decir, en efecto,

que la probabilidad que se atribuye a *a* dada la información *b* no distará mucho de 0.92; es decir, $p\,(a,b) = 0.92$. Así pues, un enunciado *b* puede hacer probable a otro enunciado (esto se llama a veces «la regla simple de inducción»)[4].

Este argumento es bastante convincente intuitivamente pero no justifica en sí mismo la afirmación de que el cálculo de probabilidades puede interpretarse del modo indicado, como una generalización de la lógica proposicional (o lógica deductiva); porque nuestro argumento intuitivo no asegura que no vaya a haber en algún punto un choque entre las leyes del cálculo formal y esas ideas intuitivas. Una justificación completa se ofrece, sin embargo, en el Apéndice *V de *L. I. C.*, donde el álgebra booleana (y, por tanto, la teoría de la composición de enunciados, es decir, la lógica proposicional) se deriva de los axiomas del cálculo de probabilidades. Esto muestra que una interpretación lógica del cálculo de probabilidades es admisible.

Si interpretamos el cálculo en el sentido lógico, entonces todos sus teoremas obtienen un status similar al de los teoremas del cálculo de proposiciones. Podemos llamarlos, por tanto, «analíticos» o «tautológicos». Aunque podemos no saber de muchas fórmulas probabilísticas si son verdaderas o falsas, todas ellas serán tautológicas de hecho, o contradictorias. Conservemos *a* y *b* como nuestro último ejemplo (el del 92 por ciento) y preguntemos el valor de $p\,(b,a)$ en vez del valor de $p\,(a,b)$. Podemos ser incapaces de responder a esta pregunta; pero la respuesta correcta será, de todos modos, tautológica y todas las demás respuestas, contradictorias. Con esto vemos que la interpretación lógica es muy diferente de la interpretación frecuencial y de la propensivista, ya que ambas interpretan la fórmula «$p\,(a,b)$», en general, como *una hipótesis que puede contrastarse experimentalmente.*

Además, tomemos algunos enunciados fijados *a* y *b* para los que obtenemos

$$p\,(a,b) = r$$

en la interpretación lógica. El valor de $p\,(a,b)$ para los mismos enunciados *a* y *b* diferirá, en general, de *r* en la interpretación

[4] [*(Añadido en 1980) La regla simple de la inducción, como lo ha mostrado David Miller en 1965, es paradójica. Ha habido una serie de réplicas a Miller, pero, que yo sepa, ninguna de ellas es sostenible y a todas ellas replicó, a su vez, Miller que, evidentemente, llevó la mejor parte en el debate. No obstante, un número bastante considerable de autores sobre este tema han continuado como si nada hubiera ocurrido. Véase David Miller, «A Paradox of Information», *The British Journal for the Pilosophy of Science* 17, 1966, pp. 59-61; y mi «A Coment on Miller's New Paradox of Information», en el mismo número, pp. 61-69.]

frecuencial o en la propensivista. Esto es evidente, en primer lugar, porque al escoger un cierto valor de *p (a,b)* —por ejemplo el valor *s*— estamos haciendo una *estimación hipotética* y, obviamente, somos libres de conjeturar la estimación que deseemos. (Desde luego, después deberíamos *contrastarla*.) Pero aparte de nuestra libertad de elección, no hay la mínima razón para creer que el *r* de la teoría lógica vaya a representar una *buena* estimación desde el punto de vista de la interpretación frecuencial o de la propensivista.

Para mostrar esto, sea *a* el enunciado (función) «*x* es una elipse» y *b* el enunciado (función) «*x* es una órbita planetaria». Entonces podemos (si queremos expresarla probabilísticamente) expresar la primera ley de Kepler —una conjetura audaz— con la expresión probabilística

$$p\ (a,b) = 1$$

puesto que esto significará: dado que *x* es una órbita planetaria, la probabilidad de que *x* sea una elipse es igual a 1, lo que equivale a decir «el 100 por cien de todas las órbitas planetarias son elipses». Por otro lado, es bastante claro que, desde el punto de vista de la interpretación lógica, *p (a,b)* será o bien igual a cero (considerando la infinidad de posibilidades, es decir, la infinidad de formas posibles de las órbitas planetarias) o casi cero. E incluso si incorporamos a *b* la evidencia proporcionada por Tycho, sobre la que trabajaba Kepler, la situación seguirá siendo esencialmente la misma. No sólo porque esta evidencia se refería principalmente a un solo planeta; sino, sobre todo, porque era fragmentaria incluso con respecto a ese planeta.

Este ejemplo muestra que una estimación hipotética muy «buena» de la probabilidad puede diferir ampliamente del valor de la correspondiente probabilidad lógica.

6. COMPARACION DE LAS INTERPRETACIONES OBJETIVAS Y SUBJETIVAS

Nuestro principal interés, si el lector me permite que se lo recuerde, es el status de las hipótesis probabilísticas en la ciencia, especialmente en física. La interpretación lógica no entra aquí, simplemente porque considera a los enunciados de probabilidades analíticos o tautológicos y, por tanto, *incontrastables*. La interpretación lógica entra, sin embargo, de forma indirecta, a través de las interpretaciones subjetivas.

Como ya se ha explicado, llamo «*objetivas*» a la interpretación frecuencial y a la propensivista: ambas consideran que las

probabilidades son propiedades de ciertos sistemas físicos, de las disposiciones experimentales, por ejemplo. Llamo «*subjetivas*» a aquellas interpretaciones que consideran que esas mismas probabilidades son medidas de nuestro propio conocimiento imperfecto respecto a lo que ocurre en esos sistemas físicos.

La interpretación subjetiva es consecuencia, principalmente, de la plausible y ampliamente aceptada concepción según la cual siempre que la probabilidad forma parte de nuestras consideraciones, esto se debe a la imperfección de nuestro conocimiento: si nuestro conocimiento fuese perfecto, no necesitaríamos la probabilidad, porque siempre tendríamos certeza.

Esta concepción es profundamente errónea. Su objetivo es dilucidar y analizar el significado de los enunciados de probabilidades en física y en la teoría de los juegos de azar. Los enunciados tales como «la probabilidad de que salga cara es igual a 1/2» se interpretan como significando «el enunciado según el cual saldrá cara y el enunciado alternativo de que no saldrá cara al tirar una moneda están, ambos, igualmente bien apoyados por la insuficiente evidencia que constituye nuestro conocimiento». En otras palabras, la interpretación subjetiva trata de reemplazar las interpretaciones que he llamado «objetivas» por otra que se cree que tiene ciertas ventajas. No obstante, la interpretación subjetiva choca con algunas de las consecuencias más sencillas, más importantes y mejor contrastadas de la teoría objetiva.

No tengo ninguna objeción, naturalmente, a que una teoría objetiva se traduzca a un lenguaje subjetivo. Si alguien prefiere decir «tengo razones para creer que el sol brilla», «tengo conocimiento de que es de día» (etc.), en lugar de «el sol brilla» y «es de día», no perderé el tiempo tratando de convertirle. Del mismo modo, si alguien quiere sustituir un enunciado objetivo de la forma «$p\ (a,b) = 1/3$» por algo parecido a: «si yo tuviera que apostar con *a*, estaría dispuesto, en vista de mi conocimiento de las *condiciones b*, a apostar 2:1 en contra suya», tampoco discutiré con él. Una reinterpretación subjetivista de este tipo, que reflejase la teoría objetiva, no sería muy interesante y no nos iluminaría mucho, pero sería inofensiva.

Sin embargo, no es la teoría objetiva —en la que *b* representa las *condiciones* (experimentales)— la que la teoría subjetiva intenta reflejar, sino la *interpretación lógica*, de la que ya hemos mostrado sucintamente que en ocasiones llevará a valores de la probabilidad que difieran de los obtenidos por la teoría objetiva.

Así pues, la posición de la teoría subjetiva es insostenible: *a*) su objetivo es analizar y explicar las probabilidades objetivas;

b) se ve obligada a tomar como base la teoría lógica (como mostraré en esta sección); pero hemos visto *c*) que las probabilidades lógicas y las objetivas pueden dar lugar a valores en conflicto, debido al hecho de que los enunciados de probabilidades lógicas son *tautologías inconstrastables* (o algo de este tipo) mientras que las probabilidades objetivas son *hipótesis contrastables* libremente escogidas.

En primer lugar, sólo mostraré *b*): que la teoría subjetiva está obligada a basarse en la interpretación lógica.

Esto ocurre porque la teoría subjetiva trata de interpretar *p (a,b)* como el grado de creencia en *a* que puede ser razonablemente justificado con nuestro conocimiento efectivo total *b*. Pero, en realidad, éste es el mismo problema (sólo que expresado en lenguaje subjetivo) que trata de resolver la interpretación lógica; porque la interpretación lógica trata de evaluar el grado en el que el enunciado *a* está respaldado, apoyado o justificado lógicamente por el enunciado *b*. Así, si interpretamos subjetivamente la interpretación lógica —esto es, en términos de nuestro conocimiento o nuestra nesciencia—, entonces *p (a,b)* es, precisamente, el grado en el que nuestro conocimiento efectivo total *b* justifica racionalmente un *a* hipotético o dudoso.

7. LAS INTERPRETACIONES OBJETIVISTAS Y SUBJETIVISTAS DE «*B*» EN «*P (A,B)*»

Mi último párrafo señala una diferencia pequeña pero significativa entre los dos enfoques. El subjetivista considera a *a* como su hipótesis y *p (a,b)* como nuestro grado de creencia en ella, mientras que el objetivista toma «*p (a,b)* = *r*» como su hipótesis. (Puede creer en ella o no.) Volveremos sobre este punto más adelante. De momento, quiero resaltar la diferencia entre las concepciones objetivistas y subjetivistas respecto al papel de «*b*» en «*p (a,b)*».

La concepción objetivista de «*b*» es que «*b*» enuncia las condiciones repetibles de una situación repetible. Como hemos visto, los resultados de experimentos previos no forman parte de la información *b*, y si formasen parte de ella, podrían omitirse sin prejuicio; porque tienen que ser no pertinentes, puesto que suponemos la repetibilidad de las condiciones *b*; y la repetibilidad entraña independencia de los predecesores.

La concepción subjetivista de *b* es muy diferente. Aquí *b* debe contener todo nuestro conocimiento pertinente; y la observación de los resultados anteriores del experimento en cuestión es muy pertinente. Sólo a partir de ellos, según el sub-

jetivista, podemos juzgar el valor de *p (a,b)*. Si nuestro conocimiento pasado —es decir, *b*— nos dice que *a* ha ocurrido a menudo en el pasado, entonces *p (a,b)* tendrá un valor mucho mayor que si *b* nos dice que *a* ha ocurrido muy raramente.

Esta plausible concepción que, de hecho, es la intuición fundamental en que se basa el enfoque subjetivo, será desarrollada en la próxima sección. Aquí sólo quiero mostrar que

i) *b* en la teoría objetiva no es, en general, lo mismo que *b* en la teoría subjetiva,

ii) incluso cuando es lo mismo o aproximadamente lo mismo, la *b* de la teoría subjetiva contendrá información que, según la teoría subjetiva debe ser sumamente pertinente, aunque debe ser no pertienente según la teoría objetiva.

Respecto a i), tanto la teoría objetiva como la lógica pueden escoger su *b* con entera libertad. La teoría lógica está interesada en *p (a,b)* para cualquier *a* y para cualquier *b*. Del mismo modo, la teoría objetiva puede plantear la cuestión de la probabilidad de *a*, dado cualquier conjunto de condiciones *b*. Naturalmente, la elección efectiva de *a* y *b* para la teoría objetiva dependerá en gran medida de la situación científica del problema. Pero no hay razón alguna para restringir a *b* a, digamos, situaciones ya observadas o conocidas de otro modo.

Contrariamente a esto, la teoría subjetiva considera a *p (a,b)* como el grado de nuestra creencia en *a* si *b* representa nuestro conocimiento efectivo total. En todas las aplicaciones, la interpretación subjetiva interpretará *b* como la suma total de lo que «nosotros» sabemos. En todas las aplicaciones, la teoría subjetiva no es, pues, libre en la elección de *b*: no es una constante porque nuestro conocimiento está siempre cambiando, pero tampoco es una variable en el sentido de que podamos escogerla libremente; por el contrario, es algo que está totalmente fuera de nuestro control: nuestro conocimiento es lo que nos es «dado».

Contrariamente a esto, la teoría objetiva considera a *b*, incluso en las aplicaciones, como abierta a la libre elección. Soy yo quien puede preguntar libremente cuál será la probabilidad de que salga cara si doblo la moneda de una determinada forma, de modo que la cara esté en el lado cóncavo. Cómo hallar la nueva probabilidad es otra cuestión. Las condiciones no están dadas, nosotros las escogemos o seleccionamos o construimos cuando escogemos nuestro problema.

Por tanto, no hay razón alguna para que, en cualquier problema real en el que se trate de determinar la probabilidad de *a*, las dos teorías tengan que operar con un mismo *b*.

Pero, contestará el subjetivista, usted es un ingenuo y una víctima de formulaciones superficiales. Si usted dice que

«escoge» *b*, sólo quiere decir que puede escoger doblar una moneda. ¿Quién lo niega? Si usted quiere hallar la probabilidad de que salga cara en el caso de una moneda doblada, sencillamente tendrá usted que utilizar la experiencia pasada si quiere obtener un resultado razonable: utilizará usted todo su conocimiento y el hecho de haber doblado precisamente esta moneda será una parte importante y pertinente de él. Verlo como el único componente de *b* es una ingenuidad. Puede ser la parte más conspicua de *b* —tan conspicua como para hecerles olvidar a usted y a otros como usted el resto y darlo por sentado—. Pero el lógico debe buscar siempre las premisas que se suponen implícitamente en una argumentación. Y esas premisas suprimidas son, en este caso, todo nuestro conocimiento pasado. No todo él será pertinente, sin duda. Pero experiencias previas con tiradas de monedas y también con la influencia de las asimetrías en las frecuencias de monedas y dados y, de hecho, todo nuestro conocimiento de la física puede ser pertinente. Y ¿cómo hallará usted el valor de *p (a,b)* —que salga cara con esta moneda— si no es observando la frecuencia relativa en una serie de tiradas? Después de cien tiradas o de unos cientos de ellas, puede usted ser capaz de calcular *p (a,b)*.

He tratado de presentar convincentemente el caso del defensor de la teoría subjetiva; no obstante, lo que él dice aquí está todo confundido y embrollado, pienso yo. Pero sólo atacaré aquí aquéllos de sus argumentos que son relativos al problema del status de *b*.

Primero, debo hacer notar que el subjetivista está interesado, igual que yo, en el problema de determinar efectivamente la probabilidad de que salga cara al tirar precisamente esta moneda doblada. En esto estamos de acuerdo; éste es terreno común, la base firme de nuestra discusión.

Ahora nuestro subjetivista afirma que el valor verdadero, mejor o adecuado, de *p (a,b)* estará fuertemente influenciado por los resultados de los experimentos anteriores.

Aquí tendremos que pararnos a pensar. La teoría objetiva dice que los resultados de los experimentos anteriores no son pertinentes: en caso contrario, los experimentos no serían una sucesión de repeticiones. El subjetivista dice que son pertinentes. ¿Quién tiene razón? Los sentimientos tienden, supongo, a apoyar al subjetivista en esto. Dejaré la respuesta para las secciones siguientes y, en cambio, preguntaré ahora: ¿es ésta simplemente una diferencia de opinión o hay algo más detrás? La no pertinencia y la independencia no son, después de todo, juguetes filosóficos o epistemológicos, sino conceptos que ambas partes definen mediante las mismas fórmulas matemáticas, bien conocidas (la ley especial de la multiplicación).

El objetivista puede señalar que su teoría de la independencia o no pertinencia es una de las partes más importantes de las aplicaciones del cálculo. Es equivalente, por ejemplo, al principio de exclusión del sistema de juego (de la que se trata en *L. I. C.*, secciones 49 y ss.), que está bien contrastada estadísticamente. la independencia puede contrastarse y ha sido contrastada y las contrastaciones son suficientemente sensibles como para detectar, por ejemplo, que los métodos habituales de barajar las cartas no son muy buenos[5].

Ahora bien, el terreno común entre la teoría objetiva y la subjetiva era que ambas querían explicar la teoría de la probabilidad tal como se usa en física o en los juegos de azar; y la teoría subjetiva afirma que puede hacer esto tan bien o mejor que cualquier teoría objetiva. Pero aquí hay un conflicto entre la teoría subjetiva y todas las aplicaciones: éstas requiren independencia de experimentos anteriores y que éstos no sean pertinentes donde la teoría subjetiva (sin duda muy plausiblemente) exige dependencia y pertinencia.

Así, el objetivista puede decir que incluso aunque adopte explícitamente las premisas supuestamente suprimidas (los resultados de los experimentos anteriores) como partes componentes de *b*, su teoría matemática le dice que puede ignorar, de todos modos, esos componentes como no pertinentes. El, por tanto, no descuida su experiencia pasada ingenuamente, sino que lo hace siguiendo las exigencias de una teoría de mucho más éxito, en realidad, la misma que el subjetivista se propone explicar.

La diferencia respecto a *b* permanece porque el subjetivista cree que hay partes de *b* que son muy pertinentes y el objetivista mantiene que hay que descartarlas por no pertinentes; sea el objetivista o el subjetivista el que tenga la razón de hecho, sus teorías son completamente diferentes y la afirmación del subjetivista de que entiende e interpreta la teoría tal como se emplea en física debe descartarse, simplemente a causa del status de *b*.

En resumen: el objetivista mantiene que, bajo condiciones físicas objetivas constantes, *b* (proporción constante entre bolas blancas y negras en la urna; mezcla completa constante) la probabilidad de cada tirada de la urna permanece constante. El subjetivista está en desacuerdo: la probabilidad cambia después de cada tirada de cuyo resultado somos informados, incluso aunque se nos informe también de la constancia de las condiciones físicas.

---

[5] Véase W. Feller, *Introduction to Probability Theory*, 1950, p. 336, esp. nota 9; véase la tercera edición, 1968, pp. 406 s., esp. nota 16.

Nota histórica: Keynes tiene el gran mérito de haber notado (*A Treatise on Probability*, 1921, pp. 342 s.) que la aplicabilidad de algunas de las partes esenciales de la doctrina clásica (como la ley del binomio y la ley de Bernoulli) no pueden ni explicarse ni sostenerse siquiera desde el punto de vista de su teoría subjetiva. Porque esta aplicabilidad supone independencia o no pertinencia, una condición que difícilmente podría obtenerse desde el punto de vista de su teoría subjetiva. Los subjetivistas posteriores, por regla general, han olvidado este punto y han operado con la teoría clásica de la independencia como si no presentase grandes problemas dentro de la teoría de ellos. Una excepción es Carnap, que respalda la advertencia de Keynes (*Logical Foundations of Probability*, 1940, p. 499) y cuyo subjetivismo es, quizá, el más puro de todos. Aunque hace distinción entre la teoría subjetiva sobre grados de creencia racional («probabilidad$_1$») y la teoría objetiva de la frecuencia relativa («probabilidad$_2$»), explica las estimaciones de frecuencia objetiva relativa como determinadas por la probabilidad subjetiva (considera, de hecho, que la estimación óptima del valor de la probabilidad objetiva es idéntica al valor de la probabilidad subjetiva); y esto nos interesa aquí: Carnap sostiene que nuestro *b* (él lo llama «*e*» o «evidencia») de la probabilidad objetiva es siempre el mismo conocimiento total que entra en las fórmulas subjetivas. Respecto a la ley del binomio, Carnap dice también, como Keynes, que «algunos de sus usos tradicionales no son admisibles» (*loc. cit.*). Pero esos son, precisamente, los usos a los que me he referido más arriba como característicos de la teoría de los juegos de azar y del principio de exclusión de los sistemas de juego. En consecuencia, la ley del binomio es también inaplicable a aquellos casos en los que Carnap afirma su aplicabilidad (véase *loc. cit*, p. 500, «el experimento descrito con la sustitución de cada bola») porque nuestra *b* —es decir, la *e* de Carnap— cambia constantemente con cada nueva tirada, puesto que cada resultado cambia nuestra experiencia pertinente. Y esto tiene que ser así incluso aunque nuestra evidencia implique información definitiva respecto a la constitución de la urna.

CAPITULO II

CRITICA DE LA INDUCCION PROBABILISTICA

8. LA REGLA INDUCTIVA SIMPLE

La interpretación subjetiva no logra analizar ni explicar la teoría tradicional de los juegos de azar. Fracasa en el análisis de sucesiones fortuitas o independientes. Tampoco se puede explicar la aplicabilidad de esta teoría. Por el contrario, las aserciones de la teoría subjetiva no pueden dejar de chocar con los resultados clásicos de Bernoulli y sus sucesores. Así, los subjetivista están equivocados respecto a un punto muy importante: al analizar la teoría clásica, muestran con su programa que no la entienden.

A pesar de lo grave que es esta crítica, no contesta las preguntas ¿Quién tiene razón? ¿Quién tiene la teoría correcta, el objetivista o el subjetivista? «Correcta» significaría aquí la teoría aplicable a la realidad física, por ejemplo, a los juegos de azar. Es concebible que la teoría clásica esté equivocada, a pesar de sus buenos resultados, y que la teoría subjetiva ofrezca una aproximación mejor.

Yo no creo que sea así. Al contrario, creo que la interpretación subjetiva es totalmente errónea, que es una mala interpretación obvia.

Pero la interpretación subjetiva es extremadamente influyente entre los físicos; no tanto en cuanto teoría bien definida y coherente como en cuanto segunda línea conveniente de defensa a la que retirarse en caso de que surja cualquier dificultad en otro lugar. De este modo ha surgido una especie de lenguaje ambiguo en la física, tanto en la mecánica estadística como en la teoría cuántica; y, a veces, encontramos que se «explican» hechos físicos objetivos recurriendo a nuestra falta de conocimiento[1].

En vista de que ésta es la situación, parece importante considerar seriamente el razonamiento que lleva a la interpretación subjetiva. Los principales argumentos son, sin duda, éstos:

La idea de probabilidad entra en nuestras consideraciones

[1] [Véase *Teoría cuántica y el cisma en física*, Tomo III del *Post Scriptum*. *Ed.*]

cuando no sabemos lo suficiente para estar seguros: si supiéramos si iban a salir caras o cruces, no necesitaríamos hablar de la «probabilidad» de caras o cruces.

La ignorancia, sin embargo, no basta. Si comparamos echar suertes con cartas y jugar con un dado, encontramos que *nuestro conocimiento* de que ahora hay seis posibilidades en lugar de dos influencia las probabilidades. Y, aún más, si consideramos una apuesta con monedas posiblemente dobladas o con dados cargados, entonces las probabilidades que supongamos dependerán en gran medida de los resultados pasados. Es difícil calcular las probabilidades para un dado cargado, pero es fácil hacerlo si tenemos tiempo de tirar el dado cien mil veces, pongamos por caso: todos admitirán que las frecuencias relativas que se obtengan en una sucesión tan larga serán estimaciones excelentes de las probabilidades.

De este modo, nuestra experiencia pasada, nuestro conocimiento de sucesos pasados determina claramente nuestras estimaciones en este caso; y hay muchas razones para creer que siempre juega un papel poderoso en la determinación de las mejores estimaciones de las probabilidades en cuestión. ¿De qué otra manera podríamos determinar las probabilidades, si no fuera a la luz de nuestra experiencia pasada, sobre todo en casos no simétricos, como el de los dados cargados? ¿Y cómo sabríamos que un dado no está cargado, si no es a la luz de la experiencia pasada?

Estas consideraciones sugieren la siguiente *regla inductiva simple:* si en un número muy grande de repeticiones de un experimento encontramos que el resultado *a* ocurre con una frecuencia *m/n*, entonces la mejor estimación de la frecuencia de *a*, con respecto a esta evidencia experimental, es igual o aproximadamente igual a *m/n*.

He dado a propósito a esta fórmula una forma que no es muy definida y, por tanto, no muy fuerte. En la forma dada sólo es aplicable a números grandes de repeticiones; y no afirma que *m/n* sea la mejor estimación: sólo afirma que la mejor estimación será aproximadamente igual a *m/n*.

Puesto que en esta forma la fórmula es débil, todas sus variantes más fuertes serán destruidas con ella si podemos refutarla, como, de hecho, podemos[2].

Porque la regla inductiva simple es falsa. Las razones se han indicado ya en la sección anterior. Un sencillo ejemplo mostrará que puede llevar a unos resultados totalmente absurdos

[2] Las variantes refutadas incluyen incontables cantidades de teorías de la inducción (entre ellas, la totalidad del llamado «*Continuo de los métodos inductivos*» de Carnap, de 1952).

cuando no se satisface la condición de *independencia* (en el sentido de la teoría objetiva). Por otro lado, la teoría subjetiva-inductiva es incapaz de combinar su «regla simple» con la exigencia de independencia. La regla misma hace que los resultados de experimentos pasados afecten a las probabilidades futuras. Estos resultados están lejos de no ser pertinentes: la regla los hace aparecer como la información más pertinente de que se dispone, sobre todo en el caso de sucesiones largas. De modo que el subjetivista no puede exigir independencia.

Un ejemplo llamativo con el que podemos refutar la regla inductiva simple es el juego «rojo o azul».

Un jugador juega a caras o cruces a crédito con su banquero: se lleva un libro en el que se le apunta un chelín en su haber cada vez que gana y uno en su debe cada vez que pierde. Si, después de cualquier tirada, el jugador debe dinero al banco, decimos que hemos observado un suceso «rojo». Si no debe ningún dinero, decimos que el suceso observado es «azul». El juego que llamo «azul o rojo» consiste en apostar sobre la ocurrencia de uno de esos dos sucesos observables, «rojo» o «azul».

Ahora bien, desde el punto de vista de la *teoría objetiva* es bastante claro que:

i) La probabilidad de rojo = la probabilidad de azul = 1/2, o casi. (Hay una ligera pero despreciable asimetría en las condiciones, en la medida en que «rojo» significa «debe», mientras que «azul» significa «haber» o «saldo cero».) Por tanto, desde el punto de vista objetivo podemos decir que

$$p\,(a,b) = (\bar{a},b) = 1/2$$

donde a es rojo, $\bar{a}$ es azul y b las condiciones del juego (de la disposición experimental).

ii) La sucesión de a y $\bar{a}$ no es independiente. Los cálculos[3] muestran que este hecho lleva a resultados muy inesperados. Si disponemos un experimento colosal, tirando la moneda cada segundo durante un año entero, entonces se cumple lo que sigue: habrá una probabilidad de aproximadamente 0,9 de que *la diferencia* entre las dos frecuencias observadas —la de a y la de $\bar{a}$— exceda 1/6 (es decir, 2 meses). Habrá una probabilidad de más de 0,5 de que la diferencia entre las dos frecuencias observadas exceda 2/3 (es decir, 8 meses). Así pues, habrá más probabilidades favorables que desfavorables de que las frecuen-

[3] Todos los resultados están tomados de W. Feller, *An Introduction to Probability Theory and its Applications*, tercera edición, capítulo III, sección 4, esp. pp. 82 s. Los artículos originales están citados allí en las notas de p. 252.

cias observadas difieran como 1/6 y 5/6, mientras que las probabilidades reales son 1/2 y 1/2.

En otras palabras, en este juego es, objetivamente, sumamente improbable que las estimaciones de acuerdo con la regla inductiva simple obtentan buenos resultados y muy probable que las desviaciones entre la estimación inductiva y el valor 1/2 sean muy grandes; y esto será así incluso aunque hagamos un número enorme de observaciones, más de las que pueden observarse en una vida.

Aquí tenemos, pues, un caso en el que la regla inductiva simple fallará casi con toda seguridad, a pesar de su plausibilidad. La razón es que los resultados sucesivos del juego «azul o rojo» no son independientes y que es probable que la regla inductiva simple lleve a resultados que concuerden con buenas estimaciones objetivas sólo en los casos de observaciones *objetivamente independientes,* cuando podemos aplicar los métodos de Bernoulli. Sin embargo, estas mismas explicaciones pertenecen esencialmente a la teoría objetiva. Es imposible traducirlas al lenguaje de la teoría subjetiva o inductiva porque, al traducirlas, la independencia objetiva se transformaría en no pertinencia subjetiva. Y no sólo es incompatible la no pertenencia subjetiva con la regla inductiva simple, sino que el carácter de no pertinencia de las observaciones en «rojo o azul» es, desde el punto de vista subjetivo, exactamente el mismo que en «cara o cruz». Desde el punto de vista objetivo, las repetidas observaciones de «rojo» o de «azul» *no son, en ningún sentido, repeticiones de experimentos:* el efecto secundario muestra que las condiciones no se reproducen. (Para repetir el experimento tenemos que iniciar un juego de «rojo o azul» de una longitud definida, desde el principio.) Pero la teoría subjetiva no tiene medios para distinguir de este modo entre repetición de observaciones y repetición de experimentos, ya que esta distinción fundamental se basa en una teoría de la independencia que ella no puede formular.

El fracaso de la regla inductiva simple en el juego «rojo o azul» puede ilustrarse comparando el comportamiento de un *jugador inductivo* con el de un *jugador racional.* Supongamos que, hasta el momento, las observaciones muestran 100.000 ocurrencias de «rojo» y 10 de «azul». Entonces el jugador inductivo estará dispuesto a apostar 10.000 contra uno por el rojo, cualquiera que haya sido la última ocurrencia. El jugador racional (que toma en consideración no sólo los resultados observados del juego, sino también su estructura matemática) sabrá, sin embargo, que en vista del efecto secundario inherente al juego, una apuesta de 1:1 por el azul será muy ventajosa para él, siempre que la última ocurrencia haya sido «azul».

El juego de «rojo o azul» no es, naturalmente, más que un ejemplo especialmente llamativo del fracaso de la regla inductiva simple si se utiliza con el fin de evaluar una aproximación razonablemente buena. Si se usa con ese fin, funciona bastante bien, desde luego, en el juego normal de echar a suertes con una moneda, que es independiente. Pero funcionaría muy mal, incluso en el caso de un juego de tirar una moneda, si se utilizase como *sistema de juego*. (Las probabilidades inducidas se desviarían en ocasiones de 1/2, de tal forma que, en esos casos, las probabilidades objetivamente desfavorables serían aceptables para el jugador.)

Resumiendo: la especiosa plausibilidad o «evidencia manifiesta» de la regla inductiva simple no es más que una mala interpretación de los ejemplos objetivamente independientes que suelen considerarse. Y la diferencia que vimos que se producía entre los ejemplos implica la idea de independencia objetiva y transciende, por tanto, la teoría subjetiva-inductiva[4].

9. COMO INTERPRETAR LA REGLA INDUCTIVA SIMPLE DONDE FUNCIONA

No hay duda de que «la regla inductiva simple» funciona a veces: si encontramos, en unos cuantos miles de tiradas con un dado, que el lado «6» aparece con una frecuencia de, digamos, un poco más de 1/4, en vez de 1/6, entonces (podría argüirse), aplicaremos la regla inductiva simple: diremos que la probabilidad está más cerca de 1/4 que de 1/6.

Estoy dispuesto a admitir esto, hasta cierto punto; pero hay que interpretarlo.

i) En primer lugar, no debemos hacer este tipo de inferencia a menos que tengamos buenas razones para creer que nues-

[4] Tomemos una sucesión tal como, por ejemplo,

0011000011110000000011111111100000000000000001111111111111111...

en la que 2^n ceros van seguidos por 2^n unos, seguidos por 2^{n+1} ceros..., de modo que las frecuencias de los ceros y los unos tienen el límite 1/2. La probabilidad de dar, accidentalmente, en esta sucesión, con un segmento aproximadamente representativo se acerca a cero para cualquier longitud elegida del segmento (incluso aunque se exija un grado muy reducido de aproximación). Esto me parece una clara refutación de la aserción de Carnap (*Logical Foundations of Probability*, p. 500) de que el teorema del binomio se cumple exactamente, *sin condición de independencia*, «del mismo modo para una *población infita*... respecto a un orden serial dado de los elementos» de esta población. [*Debo repetir aquí que, en 1965, David Miller mostró que la regla inductiva simple era lógicamente paradójica. Véase su «A Paradox of Information», *British Journal for the Philosophy of Science* 17, 1966, pp. 59-61.]

tra sucesión es *independiente*. Si surge la situación correspondiente en un juego de cartas, lo primero que debemos hacer, probablemente, es sospechar de cómo se han barajado. Pero si la hipótesis de independencia precede lógicamente a la aplicación de la regla inductiva simple, entonces no puede basarse, a su vez, en la regla inductiva: debe ser una auténtica conjetura, hecha en vista de una evaluación de la totalidad de la situación lógica; quizá sea contrastable pero, desde luego, no será una consecuencia de la regla inductiva simple.

ii) Bajo el supuesto de independencia, podemos utilizar el argumento de Bernoulli, según el cual, la mayoría de las muestras grandes (o de las sucesiones largas) será representativa. Por tanto, consideraremos que nuestra sucesión de tiradas con el dado en cuestión es representativa y concluiremos que las probabilidades de sus seis lados no son iguales y, por tanto, no simétricas: conjeturaremos que el dado esta cargado.

iii) Sin embargo, si estamos realmente interesados en el caso, comprobaremos de diversas maneras nuestra nueva hipótesis de que el dado está cargado. Una forma más de hacerlo es realizar nuevas contrastaciones estadísticas. Pero también trataremos de encontrar la posición de su centro de gravedad por medio de investigaciones físicas directas (incluido, quizá, un examen con rayos X, etc.)

iv) Si en esta investigación nos encontrásemos con que el dado es, físicamente, sumamente homogéneo y simétrico, mientras que las estadísticas continúan dando los mismos resultados, nos hallaríamos ante un enigma. En ese caso, tenemos una elección: podemos considerar que nuestras estadísticas o nuestras mediciones físicas directas están «influenciadas por un error sistemático desconocido» (todo experimentador sabe que esos «efectos ocultos» ocurren; *cf. L. I. C.*, sección 8); o si debemos atribuir al azar las desviaciones estadísticas de la simetría.

v) Si, estimulados por nuestras estadísticas, descubrimos que el dado está físicamente cargado con un pedazo de plomo, entonces sustituiremos nuestro hipotético *b* («este dado es homogéneo») por un nuevo *b* («este dado está cargado hasta tal y tal punto»). Es decir, nuestra estadística puede influenciarnos para *revisar* nuestro «conocimiento» de las condiciones. Puede hacer que nos percatemos de que las condiciones *b*, en las que creíamos como si fuesen seguras, estaban en realidad equivocadas. *Pero nuestras estadísticas no pasan, ellas mismas, a formar parte de las condiciones,* es decir, de nuestro nuevo *b*.

Este punto es de fundamental importancia: nuestra experiencia es aquí, sin duda, pertinente a la cuestión «¿cuál es el *b* apropiado en el caso de tiradas con el dado que tenemos delante?». Pero no es pertinente, en el sentido probabilístico, a

la cuestión: «¿cuál es la probabilidad de *a* en presencia de *b*?». La teoría subjetiva no distingue entre esas dos cuestiones. El que no lo haga es una consecuencia necesaria de su subjetivismo, de considerar la probabilidad como una descripción del estado de nuestro conocimiento; una concepción expresada por Peirce, que dice: «No puedo hacer una inferencia probable válida sin tener en cuenta cualquier conocimiento que yo tenga (o, al menos, cualquier cosa que se me ocurra) en relación con la cuestión»[5]. Ciertamente, de este modo podemos obtener algo parecido a un enunciado tautológico de la probabilidad lógica de *a* a la luz de nuestro conocimiento total. Pero, como muestra nuestro argumento, esto es totalmente diferente de *a*) la probabilidad objetiva real y *b*) la mejor estimación de que disponemos de esta probabilidad objetiva real.

Lo que nuestra teoría objetiva —la teoría de Bernoulli, de hecho— nos dice, es esto: en el caso de un número muy grande de experimentos independientes (o de observaciones independientes), es sumamente probable que *a*) la probabilidad objetiva real esté muy cerca de la frecuencia observada. Por tanto, la frecuencia observada será probablemente una *b*) muy buena, es decir, una estimación muy buena de la probabilidad objetiva.

Este resultado, que explica cuándo[6] puede expresarse que funcione la regla inductiva simple, forma parte de la teoría puramente objetiva. Así, la teoría objetiva *explica* el funcionamiento de la regla; y el hecho de que la regla funcione (a veces) no es, en absoluto, un argumento en favor de la teoría subjetiva.

Pero esto, naturalmente, supone entender la seriedad de la situación. Hemos mostrado que la teoría probabilística de la inducción es incompatible con la independencia, y que la independencia es una condición para el empleo lícito de la regla inductiva simple. De este modo, el empleo lícito de la regla inductiva simple resulta ser incompatible con la teoría probabilística de la inducción.

---

[5] C. S. Peirce, «A Theory of Probable Inference», *Studies in Logic* por miembros de la Johns Hopkins University, 1883, p. 161; también en *Collecte Papers*, vol II, p. 461 (*cf.* también Carnap, *op. cit.*, p. 212).

[6] No digo que la independencia sea una condición *necesaria* para que tenga éxito la aplicación de la regla inductiva simple porque es obvio que hay casos triviales como la sucesión (no independiente) 01010101... en la que la regla inductiva simple daría una estimación correcta de la frecuencia (e incluso predicciones correctas para cada caso singular). Pero para una clase grande de casos, la independencia es una condición necesaria; e incluso la sucesión de más arriba puede reinterpretarse como una sucesión 222222... en la que «2» sustituye a «01»; y una transformación de este tipo la traería, junto con todas las sucesiones causales, a la clase de las sucesiones independientes, como indica la fórmula, «si $p\,(a) = 1$, entonces $p\,(a.b) = p\,(a)\,p\,(b)$».

10. RESUMEN DEL STATUS DE «*B*» EN «*P (A,B)*»

Fundamental para todas las interpretaciones objetivas de la probabilidad es la concepción de que podemos *explicar* ciertos efectos estadísticos (físicos o biológicos, etc.) por medio de una hipótesis probabilística de la forma

$$p\ (a,b) = r$$

y que podemos contrastar esta hipótesis mediante contrastaciones estadísticas. Aquí la tarea de *b* es enunciar las condiciones del efecto en cuestión; y esas condiciones, a veces, distan de ser manifiestas (como en el caso del dado cargado): la hipótesis de que *b* enuncia verdaderamente todas las condiciones objetivamente pertinentes necesitaría ser revisada a menudo y las contrastaciones estadísticas pueden jugar un papel importante en relación con la revisión de nuestras hipótesis. La conclusión es que el enunciado $p\ (a,b) = r$ afirma describir un estado de cosas objetivo: decir que la probabilidad de sacar «tres» con este dado es ligeramente inferior a 1/6 se interpreta como un enunciado sobre el resultado de un experimento con unas ciertas disposiciones; o como enunciado de las condiciones de un experimento y de los resultados que es posible que produzcan esas condiciones. Por tanto, «$p\ (a,b) = r$» es una hipótesis exactamente igual que otros enunciados físicos; y su carácter hipotético no tiene especial conexión con su carácter probabilístico.

Esta es la concepción que comparten las diversas interpretaciones objetivas.

Contrariamente a esto, las interpretaciones subjetivas no consideran

$$p\ (a,b) = r$$

como la hipótesis en la que están interesadas; porque consideran que «$p\ (a,b) = r$» es un enunciado analítico, más que una hipótesis empírica. Pero consideran que el enunciado

$$a$$

es la hipótesis en estudio; y la fórmula probabilística «$p\ (a,b) = r$», consideran que nos informa sobre la fiabilidad o el grado de nuestra confianza racional en la hipótesis *a*, a la luz del conocimiento *b*.

Ahora bien, si va a usarse esta interpretación en la práctica, entonces necesitamos alguna regla que nos diga que si

$$p\ (a,b) = r$$

es un enunciado analítico, y si *b* es, en efecto, nuestro cono-

cimiento acumulado, total, actual, entonces, podemos aseverar que (puesto que *b* es conocido) *para el momento presente*

$$p\,(a) = r$$

es decir, una regla que nos diga que, bajo las condiciones enunciadas, nuestra presente creencia racional en *a* será igual a *r*. Esta regla es análoga al *modus ponens* de la lógica clásica; porque ambos nos permiten afirmar (con o sin especificación de una probabilidad) un enunciado *a* absoluta o incondicionalmente, enunciado que anteriormente habíamos afirmado sólo condicionalmente —dada la condición *b*—; con tal de que *sepamos* que siempre se cumple efectivamente la condición *b*. Podemos llamar a una regla de este tipo —una que nos permita afirmar absolutamente un enunciado *a*, en vez de condicionalmente, y absorverlo así, o liberarlo de los grilletes de su condiciones— una «*regla de absolución*».

Hay dos diferencias significativas entre la regla clásica de absolución (*modus ponens*, llamada a veces regla de desatamiento) y su contrapartida probabilística: en primer lugar, el resultado de la operación de absolución es en un caso la aserción de *a* o la aserción de la verdad de *a*; en el otro caso, la aserción de que *a* tiene una cierta probabilidad (grado de creencia racional). En segundo lugar, la base para la aplicación de la regla es, en un caso, la aserción de *b* o la aserción de la verdad de *b* (donde *b* es la condición o antecedente de *a*); en el caso probabilístico, es insuficiente afirmar que *b* es verdad: debemos afirmar que, en efecto, *b* es verdad, o que sabemos que *b* es verdad; debemos afirmar también que *b* comprende *todo* lo que sabemos en el momento presente; es la suma total de nuestro conocimiento. Sólo si afirmamos que *b* es nuestro conocimiento total en el momento actual, podemos deducir cuál es la probabilidad actual de *a* o cuánto podemos confiar en *a*.

Ahora bien, mi argumento es que este análisis muestra por qué la teoría subjetiva tiene que hacer absoluta la probabilidad de *a*. En efecto, no podemos *actuar* sobre la base de una probabilidad lógica *relativa* igual que no podemos hacerlo sobre la de una información condicional: ambas necesitan que se les apliquen las respectivas reglas de absolución, antes de poder proporcionarnos el tipo de información que es una base útil para la acción práctica. Y tampoco podemos actuar sobre la base de una probabilidad *lógica* relativa, igual que no podemos hacerlo sobre una tautología. Pero el enunciado absoluto

«la probabilidad de *a* es (en el momento presente) igual a *r*»

que resulta de aplicar la regla de absolución, no es ya analítico:

es un enunciado sintético, un enunciado basado en *b*, en nuestra experiencia y que cambia con ella; porque el valor de *r* cambiará, en general, con cada cambio de *b*. *Esta es la forma en que aprendemos por experiencia; al menos según la teoría subjetiva.*

Esta es la línea que debe adoptar la teoría subjetiva. Pero esta línea hace difícil, si no imposible, aceptar la concepción de que la teoría subjetiva tiene relación con problemas tales como la probabilidad de que el Sr. Andrews sobreviva los próximos cinco años (en el sentido en que puede apostar por ello una empresa de seguros de vida) o la probabilidad de que salga cinco con este dado.

En la práctica de los *seguros* podemos distinguir con bastante claridad entre negocios de dos tipos que pueden caracterizarse como i) apuesta justa o apuesta racional e ii) juego de azar. El primer tipo se basa en los grandes números y en estadísticas racionales y la prima que se cobra es muy competitiva y por ello, razonablemente «justa». El segundo tipo es muy diferente. Es un seguro contra riesgos raros o, incluso, únicos como, por ejemplo, el de que un cierto dictador empiece una guerra dentro de un determinado período de tiempo o el de que el impuesto sobre la renta disminuya en el próximo presupuesto. Las estadísticas no ayudan aquí, porque los sucesos en cuestión son raros, si no únicos.

Ahora bien, a lo que quiero llegar es a esto: el seguro de vida ordinario, que es del tipo i) y se basa en estadísticas racionales, está muy lejos de tener en cuenta, en cada caso, todo el conocimiento pertinente del que dispone. En cambio, se basa en un método de cuestionario aproximado que tiene por objeto incluir a cada caso en una clase muy amplia. Si, por alguna razón, no puede hacerse esto, el caso se convierte en tipo ii).

Dicho de una forma diferente, se sabe lo suficiente en cada caso como para poner a la persona que quiere un seguro sola en una clase. (*Prueba:* el solicitante debe ser identificable como individuo sobre la base de su propuesta.) Y podría adquirirse cada vez más información sobre la persona: algunas personas considerarían pertinente que se supiera su horóscopo, otras sus hábitos de lectura y aun otras el hecho de si come yogur habitualmente. Pero si la compañía de seguros trata el caso como perteneciente al tipo i), entonces no tomará en cuenta demasiadas preguntas. Sólo la edad y una buena salud media se consideran fundamentales. Esto es así porque los tipos i) se suponen *repetitivos* y se supone que pertenecen a una enorme clase de repeticiones. La información que se busca tiene que definir las condiciones del experimento y, si esas condiciones convierten en único al experimento, o, incluso, en raro, entonces el

método estadístico del tipo i) se hace inaplicable. Porque hay dos condiciones básicas para aplicar métodos bernoullianos: la independencia de las repeticiones (que, como hemos visto, no pueden considerarse repeticiones si no son independientes) y los *grandes números*.

Demasiada información pertinente, si se tiene en cuenta, hará siempre que el caso sea único: sacará de la apuesta «racional» o «justa» a los casos del tipo i) y los pondrá entre los casos del juego de azar irracional del tipo ii).

De este modo, los métodos racionales y estadísticos se basan en el descuido consciente de la información pertinente de que se dispone. Esto se debe al hecho de que b se considera como definitorio de condiciones experimentales repetitivas, más que como un resumen de nuestro conocimiento pertinente total. Lo que la compañía de seguros trata de hacer es hallar un r razonablemente estable para un b no demasiado específico. Este procedimiento contrasta claramente con la teoría subjetiva, según la cual b crece constantemente y r cambia constantemente como consecuencia de ello.

Puede decirse que los casos del tipo ii) son más parecidos a lo que tiene en mente el defensor de la teoría subjetiva, porque con ellos el asegurador hará uso de la experiencia de toda una vida: todo su conocimiento sobre los hombres, la política, etc., entrará en sus cálculos. Esto es verdad y, sin embargo, los casos ii) no son ni estadísticos ni racionales. La regla inductiva simple no puede aplicarse aquí y, evidentemente, no hay método racional que pueda hacer posible basar la estimación del riesgo del asegurador en nada que se parezca a una inferencia lógica.

Si consideramos ahora los tipos i) e ii) en relación con la regla probabilística de absolución, entonces nos encontramos con que operamos con esta regla sólo en los casos de tipo ii), es decir, en los de apuesta irracional, al azar. Apostar por sucesos únicos es, desde luego, una especie de acto absoluto: una vez que nos hemos decidido (basándonos en la totalidad de nuestra experiencia), nos hemos embarcado en nuestra aventura y desde ese momento ya no tenemos en cuenta nuestra experiencia total b.

El tipo i) es muy diferente y no da lugar a la aplicación de la regla probabilística de absolución. Está claro que no tiene sentido alguno decir que, dada la verdad de la información de que el Sr. Andrews pertenece a la clase b, podemos decir absolutamente que la esperanza de vida del Sr. Andrews es igual a r. Por el contrario, sólo en la medida en que él es un miembro de la clase definida por b, su esperanza de vida es igual a r; como hemos visto por el hecho de que podemos saber que el

mismo Sr. Andrews también pertenece a las clases determinadas por *b'* (su horóscopo), *b''* (sus hábitos de lectura), etc., y que sus esperanzas de vida en cuanto miembro de esas clases es diferente (simplemente porque *cualquier* cambio en la clase de referencia puede producir un *r* diferente). Los teóricos frecuenciales (y yo mismo cuando escribí *L. I. C.*) podrían haber expresado esto diciendo que no podemos definir la probabilidad de un suceso, sólo la probabilidad de un suceso en cuanto miembro de una clase de sucesos. Hoy, yo diría que esas condiciones objetivas de las que se conjetura que caracterizan el suceso (o experimento) y *sus repeticiones* determinan la propensión y que, en la práctica podemos hablar de la propensión sólo en relación con esas condiciones repetibles seleccionadas; porque en la práctica nunca podemos, desde luego, considerar *todas* las condiciones bajo las cuales ha ocurrido un suceso real o ha tenido lugar un experimento real.

Por tanto, en cualquier hipótesis probabilística explicativa, parte de nuestra hipótesis consistirá en suponer que tenemos la lista de las condiciones pertinentes, *b*, características del tipo de suceso que queremos explicar. El subjetivista considera *b* como dado, como datos inalterables, mientras que su hipótesis es *a* como hemos visto. El objetivista, por regla general, no toma *a* como hipótesis, sino sólo el valor de *r*.

Respecto a *a*, la actitud de ambos es, una vez más, totalmente diferente. El subjetivista puede encontrar una hipótesis interesante, *a* —la hipótesis de su preferencia, podríamos decir—, y al preguntar sobre *p (a,b)* está preguntando si puede contarse con *a* (dado *b*). El objetivista, por otro lado, siempre que esté considerando la fórmula $p\,(a,b) = r$, no estará más interesado en el suceso determinado *a* que en el suceso no-*a*. No considera *a* como su hipótesis preferida, sino como uno de los sucesos posibles; y en lo que está interesado no es en el propio *a*, sino en la probabilidad de *a*. Así pues, el actuario de seguros, que es el responsable de estimar las probabilidades de vida, no está más interesado en el hecho *a* (que un cliente haya sobrevivido) que en el suceso no-*a* (que un cliente haya muerto): ambos son material equivalente para él. (Esto es así, aunque su compañía mucho más «interesada» en *a* que en no-*a*, en el sentido de que deseará sinceramente que todos sus clientes lleguen a una edad muy avanzada.)

Volviendo a la probabilidad de sacar «cinco» con un dado, podemos hacer notar otras discrepancias entre la teoría subjetiva y la objetiva y la práctica.

Según la teoría subjetiva, en un juego de azar debemos determinar (por medio de una u otra forma de la regla simple) la probabilidad subjetiva del suceso, digamos que «cinco» al

tirar un dado determinado. Debemos hacerlo, por ejemplo, observando juegos con dados en general y con un dado en particular y compilando algún tipo de estadísticas. Después de determinar de este modo el grado correcto o racional de creencia, es racional actuar basándose en él, es decir, aceptar las probabilidades correspondientes *absolutamente*. Es decir, volvemos a aplicar la regla de absolución.

Pero esto no describe, en absoluto, la actitud normal de un apostador. Este apuesta siempre *suponiendo que las condiciones* —es decir, *b*— estan satisfechas; por ejemplo, que el dado no está cargado. Ni siquiera su apuesta es absoluta: si más adelante descubre que el dado está cargado, siente que tiene derecho a cancelar la apuesta. O, si descubre que se le ha hecho otra «trampa» implícitamente prohibida por las condiciones aceptadas del juego, de nuevo considerará que no está obligado. Así pues, no apuesta por *a* (como afirma el subjetivista), sino que apuesta por *a* siempre que se satisfagan ciertas condiciones *objetivas b*.

Puede añadirse, de paso, que el subjetivista estaría en un error si creyera que podía interpretar esta situación diciendo que no apostamos que *a* va a ocurrir, sino que, más bien, apostamos por un enunciado condicional. Porque la probabilidad de un enunciado condicional es muy diferente de la probabilidad condicional (o relativa), como puede verse en lo que sigue. Sea *b*, de nuevo, nuestro conocimiento total, *c*, las condiciones del juego; entonces, puede sugerir el subjetivista, no apostamos por *a* (dado *b*) sino por «si *c*, entonces *a*» (dado *b*), y tras aplicar la regla de absolución, por «si *c*, entonces *a*», absolutamente. Esta interpretación no es compatible con las reglas del cálculo de probabilidades, puesto que «si *c*, entonces *a*» tendrá una probabilidad mayor que *a*, a menos, claro está, que *c* sea parte de *b*, en cuyo caso, la condición *c* pierde su fuerza; es decir, en ese caso tenemos $b = bc$ y, en consecuencia, p (si *c*, entonces *a*, *b*) $= p(a,b)$ y no hay razón para que, después de aplicar la regla de absolución, obtengamos la probabilidad actual de «si *c*, entonces *a*» en lugar de la probabilidad actual de *a*.

Así pues, el papel de la evidencia *b* en la teoría subjetiva es fundamentalmente diferente del papel de las condiciones experimentales *b* en la teoría objetiva; y no parece que la teoría subjetiva se ocupe, ni por lo más remoto, de los hechos y los problemas en los que está interesada la teoría objetiva.

11. LA UTILIDAD DECRECIENTE DEL APRENDIZAJE POR INDUCCION

Estrechamente relacionada con estas consideraciones está la siguiente diferencia entre las interpretaciones objetiva y subjetiva.

El partidario de la teoría objetiva conjetura que, manteniendo constantes las condiciones b, obtendrá un r bastante estable. Podrá calcular r si tiene una muestra muy grande porque Bernoulli le dice que la muestra será muy probablemente *representativa*, esto es, mostrará una frecuencia muy cercana a ese r que tratamos de determinar.

Por el contrario, el partidario de la teoría subjetiva opera con un b cambiante y, en consecuencia, con un r cambiante. Puede decir que, con la acumulación de su experiencia estadística, obtendrá, de hecho, un r cada vez más estable; e intentará identificar el r objetivista con ese r suyo, cada vez más estable.

Entre las diversas dificultades que crea esta concepción pueden destacarse las siguientes, para discutirlas.

Según la concepción subjetivista, los cambios en r reflejan con precisión *el modo en que aprendemos por experiencia*. Y el hecho de que r se haga cada vez más estable refleja el modo en que el aumento de nuestro saber estabiliza nuestras creencias racionales.

El hecho de la estabilización misma no puede ponerse en duda: según la regla inductiva simple, r será aproximadamente igual a m/n, en donde n es el número total de nuestras observaciones pertinentes y m el número de observaciones en las que se observó la misma propiedad expresada por a (o algo favorable a a); aquí, una vez más, a es la hipótesis que estamos considerando y b la evidencia a partir de la cual obtenemos las estadísticas expresadas por m/n; y tenemos

$$p\,(a,b) = r \approx m/n$$

Ahora bien, si n se hace muy grande, entonces cualquier nueva observación tendrá muy poca influencia en r, porque si n es muy grande, $(m + 1)/(n + 1)$ ó $m/(n + 1)$ será casi lo mismo que m/n.

Este simple hecho significa que r debe hacerse relativamente estable cuando hemos acumulado experiencia y n se ha hecho grande.

Pero este hecho tiene un aspecto diferente, menos favorable. Muestra que la teoría subjetiva del aprendizaje atribuye una autoridad inmensa a nuestro pasado. Después de toda una vida aprendiendo, no puede haber esperanza de aprender más: la

autoridad de nuestra experiencia pasada hace prácticamente imposible una revisión. Dicho de otro modo, cuanto más viejos nos hacemos, nos hacemos más oxidados, más lentos, atados por el pasado. Y lo mismo debe ocurrir en la ciencia.

Incluso aunque los hechos descritos fueran verdad, como cuestión de hecho psicológico e histórico (y, aunque pueden ser verdad en algunos individuos, evidentemente no lo son respecto al desarrollo histórico de la ciencia), representan claramente una teoría del aprendizaje extrañamente insatisfactoria. Es una teoría de la utilidad decreciente.

Existe algo parecido a una réplica muy obvia a este ataque crítico. Es ésta. La creciente estabilidad de nuestras creencias es, simplemente, la consecuencia del hecho de que nuestras creencias están cada vez mejor fundadas en la experiencia. Van haciéndose cada vez más fiables. Si hemos progresado en nuestro intento de saber, no debemos quejarnos si, cuanto más progresamos, más nos acercamos a algo parecido a un estado de inalterabilidad.

Mi contrarréplica a esta defensa: la verdadera inalterabilidad no se alcanza de este modo ni se la implicaba en el ataque crítico. Si aprender significa cambiar el valor *r* de la probabilidad, y si esos cambios son mejoras, entonces tiene que existir algo parecido a un valor «verdadero» de la probabilidad, un verdadero valor de *r* al que nos acercamos por medio de nuestros valores inductivos. Lo que el apologista subjetivista que me replicó tenía en mente era algo como esto: tirando muy a menudo este dado, nuestros valores de la probabilidad subjetiva se aproximarán, con el tiempo, cada vez más al valor verdadero.

Pero ¿qué es el valor verdadero de una probabilidad sino el valor objetivo, el valor del teórico objetivista? Y ¿no era, acaso, la sustitución (o eliminación) de este valor «verdadero» u «objetivo» uno de los principales fines de la teoría subjetiva, un fin escogido a causa del carácter básicamente insatisfactorio de las teorías de la probabilidad?

Apelando a algo como un valor «verdadero» u «objetivo» de la probabilidad, el subjetivismo cometería suicidio: sería como ceder el papel a otro[7].

[7] No puedo ver sino un recurso a la probabilidad objetiva en la siguientes líneas de Carnap, en «Inductive Logic and Science», (*Proceedings of the American Academy of Arts and Sciences*, vol. 80, núm. 3, marzo 1953, p. 191): «La respuesta a la pregunta: '¿cómo haremos, entonces, de larga la serie de tiradas con el dado para determinar la probabilidad?' es la misma que la respuesta a la pregunta: '¿cómo debe ser de preciso un termómetro para medir la temperatura?'. En ambos casos, la respuesta depende, por un lado, del tiempo y el dinero disponibles y, por otro, del grado de precisión que se desee. Más espe-

Pero, aparte de esto, el argumento de la utilidad decreciente no se ha respondido en absoluto. La utilidad decreciente es, como se ha mostrado, el simple resultado aritmético de la acumulación de experiencia, tanto si esta acumulación ha logrado su propósito de acercarse cada vez más a un valor «verdadero» como si no. No tenemos razón alguna para creer que la creciente lentitud del aprendizaje es una consecuencia de una mejor aproximación; en realidad, podemos ver claramente que es totalmente independiente de ella. (Recuérdese el juego de «rojo o azul».)

12. LA PARADOJA DEL APRENDIZAJE INDUCTIVO

La crítica de la teoría subjetiva-inductiva del aprendizaje nos acerca más al núcleo de la interpretación subjetiva que ninguna de mis observaciones anteriores. Porque en la obra de la mayoría de los objetivistas encontramos el siguiente argumento, al que apodaré el argumento «transcendental» a causa de su estrecha semejanza con el argumento en el que Kant apeló al *hecho* de que la ciencia pura, o el conocimiento que es válido *a priori*, existe efectivamente.

El argumento transcendental de los subjetivistas puede expresarse como sigue.

Nadie sabe mejor que nosotros los subjetivistas que la teoría de la inducción es extremadamente difícil. Por eso es por lo que trabajamos en ella. Las dificultades que nos presenta usted ni nos chocan ni nos sorprenden. Son el pan nuestro de cada día. Pueden ser muy reales y puede que no seamos capaces de resolverlas aquí y ahora. Pero *nosotros los subjetivistas sabemos una cosa: que tienen que tener solución.* Porque sabemos que aprendemos por la experiencia: esto es un hecho; y ningún sofisma hará tambalearse nuestra confianza en este hecho.

No tengo ninguna intención de negar tal hecho; y aunque tengo la suerte de no haberme sentido tan abrumado por los ataques o las dudas de los escépticos como para recurrir al argumento trascendental, admito con mucho gusto que este es un argumento que merece respeto. Tanto, que estoy dispuesto a aceptarlo aquí para volver las tornas contra los subjetivistas: trataré de mostrar que si adoptamos la teoría subjetiva-inductiva del aprendizaje, *el aprender se vuelve imposible.*

---

cíficamente, depende de las ventajas teóricas o prácticas que se espere obtener de una mayor precisión. Cuanto más preciso sea el termómetro y cuanto más larga la serie de tiradas, mayor será la precisión que se logre. En ninguno de los dos casos existe un procedimiento perfecto».

Consideremos primero lo que ocurrirá si escogemos, para el *b* de una fórmula lógico-subjetiva $p\ (a,b) = r$, aquella información que nos da precisamente las condiciones objetivas del experimento y nada más; más especialmente, si *no* incluimos, en *b* ninguna información sobre resultados anteriores (el resultado *a* o $\bar{a}$) del experimento ni sobre sus frecuencias. Suponemos que el experimento no es de los que se realizan con un objeto de jugar perfectamente simétrico (un dado o una moneda, etc.), sino que se trata de un experimento con un carácter más parecido al de los experimentos físicos normales, por ejemplo, uno que contraste la teoría estadística de los gases. Esta claro que, en este caso, el valor de *r* de la probabilidad lógica (o subjetiva)

$$p\ (a,b) = r$$

no corresponderá, en general, al valor «objetivo» *r*; si no, todo nuestro conocimiento objetivo sería lógicamente verdadero, o analítico; y no sería necesaria ninguna inducción, ni siquiera sería posible: *no podríamos aprender por experiencia.*

Obtenemos así un resultado paradójico, si suponemos que podemos identificar la *b* de la teoría subjetiva con la información sobre las condiciones objetivas. Pero obtenemos resultados igualmente paradójicos si abandonamos dicha suposición. Y, de hecho, tenemos que abandonarla; porque el *b* de la teoría subjetiva tiene que incluir conocimiento no sólo sobre las condiciones objetivas, sino también sobre *los resultados* pasados del experimento.

La teoría subjetiva-inductiva nos dice que aprendemos aplicando alguna forma de la regla simple de la inducción. Nos dice que podemos aprender, aplicando esta regla, a esperar «cara» con una probabilidad de 1/2 si vemos que se ha lanzado un penique en el pasado y que ha caído mostrando cara con una frecuencia de 1/2.

Pero esto es precisamente lo que no podemos aprender, desde el punto de vista de la teoría subjetiva. Porque es imposible para esta teoría operar con la idea de *las condiciones objetivas de un experimento.*

Como hemos visto, la teoría subjetiva tiene que interpretar *b* en

$$p\ (a,b) = r$$

como resumen de todo nuestro conocimiento, o de todo nuestro conocimiento pertinente. Esto es consecuencia de su concepción básica según la cual «$p\ (a,b) = r$» ha de interpretarse como un enunciado sobre el grado de nuestra creencia racional, *racional en vista de la evidencia en nuestra posesión.*

Así, dentro de la teoría subjetiva, *b* no puede ser un enunciado de las condiciones experimentales. Se admite que estas condiciones pueden formar parte de *b*, puesto que podemos estar informados de ellas o podemos haberlas observado. Pero como subjetivistas, no podemos separarlas del resto de nuestro conocimiento. El objetivista las separa con ayuda de su idea de pertinencia. Pero la idea de pertinencia lleva al subjetivista a un *b* muy diferente, ya que los resultados de los experimentos anteriores son de la mayor pertinencia para él, mientras que son no pertinentes para el objetivista, puesto que los nuevos resultados tienen que ser independientes de los antiguos.

Como consecuencia, el subjetivista no puede expresar la idea de *la repetición de un experimento.* Porque todo experimento procederá bajo condiciones pertinentes esencialmente diferentes (no condiciones objetivas, sino condiciones en cuanto al conocimiento). El décimo-octavo experimento procederá bajo condiciones diferentes de las del décimo-séptimo; porque entre sus condiciones estará nuestro conocimiento del resultado del décimo-séptimo experimento.

Por tanto, no podemos aplicar la regla inductiva simple, puesto que no podemos repetir nunca las condiciones de un experimento.

Podemos expresar la paradoja de este modo:

Supongamos que nuestro conocimiento aumenta, de acuerdo con la teoría subjetiva, si, y sólo si, observamos una repetición de un experimento. Entonces no puede aumentar; ya que como su aumento alteraría las condiciones *conocidas* de cualquier experimento, ningún experimento puede repetirse nunca.

En otras palabras, el supuesto de que el nuevo experimento es una repetición del antiguo es contradictorio desde el punto de vista subjetivo.

Porque si es una repetición, entonces la regla inductiva simple es aplicable, lo que convierte a todos los casos previos en condiciones sumamente pertinentes, de modo que debe ser un caso esencialmente diferente de esos casos anteriores. Así que no puede repetirse ningún experimento.

(Zenón hubiera dicho: si es el mismo, entonces no es el mismo. Luego no puede ser el mismo.)

La repetición de un experimento resulta ser, por tanto, un concepto ajeno a la teoría subjetiva. Esta teoría puede emplear, sin embargo, la idea de repetición de una *observación,* o de un hecho observado. No tiene dificultad para aseverar que hemos observado cincuenta tiradas en las que ha salido cruz. Lo que no puede aseverar es que, en esos cincuenta casos, las condiciones fueran iguales o iguales en gran medida o principalmente iguales; porque según las presuposiciones de la regla inductiva,

era, en cada caso comparado con cualquiera de los otros casos, diferentes en el sentido más pertinente.

Contra esta forma de argüir puede objetarse que he mezclado astutamente dos sentidos de la palabra «condición» y de la palabra «conocimiento». Se dirá que debemos distinguir, i) nuestro conocimiento de las condiciones objetivas del experimento e ii) las condiciones psicológicas o subjetivas existentes cuando observamos el nuevo experimento (las cuales incluyen, claro, nuestro conocimiento de resultados previos). Mi réplica es que la teoría subjetiva no puede hacer esta distinción. Todo lo que puede hacer es distinguir entre condiciones pertinentes y no pertinentes; y de la sola cosa que podemos estar seguros es de que el conocimiento de los resultados anteriores tiene que ser sumamente pertinente desde el punto de vista de la teoría subjetiva[8].

Así pues, aprender por experiencia es imposible, según la teoría subjetiva-inductiva[9]. Pero sabemos que aprendemos por experiencia.

Luego, la teoría subjetiva-inductiva tiene que ser falsa.

Este uso negativo del argumento transcendental es el único uso adecuado. No debe darse, como se hace tan a menudo, para argüir en favor de ninguna teoría sobre el aprendizaje en particular; porque es un argumento que siempre equivale a la afirmación, más o menos explícita, de que la teoría en cuestión es la única posible. (Es sorprendente ver cuántas teorías incompatibles han afirmado ser las únicas posibles.)

13. UNA MAQUINA INDUCTIVA

A pesar de mis objeciones a la teoría inductiva del aprendizaje, creo en la posibilidad de máquinas inductivas... de cierto tipo.

Para ver la posibilidad de una máquina inductiva, consideremos *un universo simplificado* que la máquina tendrá que investigar. El universo puede consistir, por ejemplo, en «indivi-

---

[8] Véase, por ejemplo, R. Carnap, *The Continuum of Inductive Methods*, 1952, p. 14, Regla C9, que afirma que «la probabilidad (en el sentido lógico) de que la próxima tirada de un dado determinado dé un número impar es *generalmente considerada (sic)* como dependiente *solamente* del número de resultados impares y del número de resultados pares obtenidos hasta el momento con ese dado». (Las cursivas son mías.)

[9] Compárese con este argumento mi análisis de las concepciones de Hume sobre la inducción, en las secciones IV y V de mi artículo «Philosophy of Science: A personal Report», reimpreso en *El desarrollo del conocimiento científico (C. & R.)*, cap. 1; y también Apéndice *X de *L. I. C.*

duos» o en «sucesos individuales» con un cierto número limitado de propiedades. Escogemos como individuos bolas de una pulgada de diámetro. Un suceso individual puede consistir en la aparición de una bola en el extremo de un tubo; luego rueda por un canalillo y desaparece por un agujero. Debe ocurrir un suceso individual de este tipo cada segundo.

Las bolas pueden tener ciertas *propiedades*. Pueden ser de acero o de cobre y pueden estar pintadas de una serie de colores diferentes. Supongamos ahora que puede hacerse funcionar a nuestro universo artificial de acuerdo con una serie de «leyes naturales». Por ejemplo, podemos arreglarlo de tal modo que sus bolas de acero entren siempre en sucesiones de tres y después de éstas, una sucesión larga pero variable de bolas de cobre; o podemos arreglarlo de modo que de cada centena de bolas de acero, 10 sean azules, 20 verdes, 30 rojas y 40 amarillas: ésta sería una «ley» estadística. Pueden disponerse otras «leyes».

Nada más respecto nuestro universo simple. Ahora, sobre la máquina de inducción. Esta puede construirse de tal modo que en un período razonable de tiempo «descubra» las leyes que se cumplen durante ese período en su universo. Y mostrará su capacidad descubriendo, si se cambian las leyes de su universo, el nuevo conjunto de leyes. (Naturalmente, habrá que dar bastante tiempo para que se realicen esos descubrimientos y las leyes, estadísticas o de otro tipo, deben mantenerse constantes durante ese tiempo.)

La máquina de inducción deberá estar equipada con un detector (quizá una aguja magnética) que le permita distinguir entre bolas de acero y bolas de cobre y otro detector de colores. Además tendrá que estar equipada con mecanismos para contar. Tendrá que ser capaz de hacer estadísticas de las diversas ocurrencias distinguibles y de calcular medias. Se la adapta aún más al *tipo* de ley (no a las leyes efectivas) que puede presentar el universo —leyes de sucesión, frecuencias generales o condicionales de una cierta estabilidad, etc.— entonces se puede construir de tal modo que sea bastante eficaz, por ejemplo en la formulación de hipótesis y en su contrastación y eliminación. Así, podrá detectar regularidades simples en el universo sin ninguna dificultad: *puede, por tanto, aprender por experiencia.*

Es útil discutir algunas de las fases más primitivas en la evolución de las máquinas de inducción con un poco más de detalle.

i) La fase más primitiva en la evolución de la máquina de inducción puede describirse como sigue:

La máquina observa las propiedades (o predicados) de cada

suceso singular. Es decir, si se trata de acero o cobre, azul o verde, etc. *Cuenta* los sucesos y también cuenta los sucesos que caen dentro de cualquiera de las diversas propiedades y los que caen dentro de dos o más propiedades («acero y azul»). Además calcula inmediatamente los respectivos cocientes, los anota con símbolos y los corrige después de cada nuevo suceso, de modo que, en cualquier momento, puede darnos las frecuencias relativas de la ocurrencia de las diversas propiedades y la combinación de propiedades.

Si hemos construido nuestro «mundo» de modo que esas frecuencias sean razonablemente estables (o si sabemos que está construido así) entonces las frecuencias calculadas por la máquina se aproximarán cada vez más, desde luego, a las frecuencias «verdaderas» u «objetivas» que hemos incorporado a nuestro «mundo»; y podemos decir entonces que la máquina puede responder en cualquier momento a una pregunta como «¿cuál es la probabilidad de que el próximo suceso posea la propiedad *P* o, quizá, el conjunto de propiedades *(P,Q,R)* con una aproximación razonable al valor objetivo?».

En la fase que acabo de describir, la máquina utiliza la regla inductiva simple (como yo la he llamado) en su forma más primitiva. Contra esta forma se ha planteado la objeción de que los primeros resultados —los de los pasos iniciales de la inducción— no son razonables. Por ejemplo, si la primera bola es de acero y azul, entonces la máquina tendrá que atribuir a los predicados «azul» y «acero» la probabilidad 1 y a los demás la probabilidad cero. La objeción planteada aquí parece de poca importancia; porque si queremos evitar tales resultados iniciales siempre podemos construir nuestra máquina de tal modo que empiece a realizar predicciones probabilísticas sólo a partir del milésimo suceso, digamos, o después de cualquier otro número *n* que elijamos, teniendo en cuenta el número de propiedades diferentes en nuestro «mundo». (El problema es tan trivial que no merece la pena hacer el esfuerzo de resolverlo sistemáticamente; porque sabemos, después de todo, que las aplicaciones de la regla inductiva simple nunca nos darán más que aproximaciones cada vez mejores)[10].

[10] En su *Logical Foundations of Probability*, 1949, y en la continuación, *The Continuum of Inductive Methods*, 1952, Carnap está interesado en construir «métodos inductivos» que correspondan estrechamente a lo que yo he descrito aquí como la fase más primitiva i) en la evolución de una máquina de inducción; y puede mostrarse que los diversos métodos inductivos contenidos en su *Continuum* son otros tantos intentos de resolver el problema que acabo de describir como trivial; es decir, el problema de cómo evitar los resultados iniciales contra-intuitivos de la regla inductiva simple en su forma más primitiva. Véase, por ejemplo, *Logical Foundations of Probability*, p. 227, donde Carnap

ii) En una segunda fase, ligeramente menos primitiva, la máquina de inducción puede, de modo similar, formular las estadísticas de pares consecutivos, ternos, etc., de sucesos. Puede descubrir, de este modo, que la probabilidad de que a un par de cobre-acero le siga un suceso de cobre o un par de acero-cobre, o cualquier terno, excepto un terno acero-acero-cobre, es cero. (Lo que significa que la máquina ha descubierto la «ley» de que los sucesos acero tienden a ocurrir, si ocurren, en una sucesión exactamente de tres.)

iii) Pueden ahora introducirse nuevas fases en su evolución, al permitir a la máquina que formule hipótesis y las contraste. No profundizaré aquí en el interesante pero difícil problema de cómo puede hacer esto una máquina. En cambio, vuelvo a la pregunta:

Si una máquina puede aprender por experiencia de esta manera, es decir, aplicando la regla inductiva simple, ¿no es obvio que nosotros podemos hacer lo mismo?

Naturalmente, nunca he dicho que no pudiéramos aprender por experiencia. Tampoco he dicho nunca que no podamos utilizar con éxito la regla inductiva simple, si las condiciones objetivas son apropiadas. Pero sí que afirmo que no podemos descubrir por inducción si las condiciones objetivas son apropiadas o no para el uso de la regla inductiva simple.

Esto puede ilustrarlo nuestra máquina en su fase i).

Si nuestra máquina obtiene buenos resultados en su fase i), *entonces es porque hemos construido nuestro «mundo» de tal forma que la regla inductiva simple puede aplicarse con éxito.* Que esto es así es obvio, si recordamos nuestro juego «rojo o azul». Nada es más fácil que construir un «mundo» que incorpore este juego o una versión generalizada de él. Pero, como

---

expone lo que él llama la «*regla directa*» que es nuestra regla simple, interpretada como una regla que da resultados exactos en vez de aproximados. (Esta es la razón de que Carnap plantee *dos* objeciones: una en relación con los resultados iniciales de la regla, y la otra en relación con el hecho de que difícilmente debemos atribuir en todo momento de tiempo una probabilidad extrema, tal como cero, a todo suceso que no haya sido observado hasta ese momento; sin embargo, esta segunda objeción desaparece si sustituimos «cero» por «no lejos de cero».) El «continuo de los métodos inductivos» de Carnap consiste, simplemente, en varios métodos para obtener resultados no excesivamente implausibles incluso por debajo del número que, en el texto, he llamado «*n*», y para elegir ese número *n*. Como él cree que la teoría de la inducción debe ser una teoría lógica, intenta dar en su *Logical Foundations of Probability* (pp. 563 ss.) una especie de justificación lógica para escoger una determinada solución de este problema. Este intento, al parecer, lo abandona en su *Continuum*, donde pone de relieve la diversidad de las posibles soluciones. (Así, el número λ de Carnap y el número κ de Kemeny son algo parecido a sustitutos sumamente sofisticados del número que yo he llamado aquí «*n*».)

hemos visto, este «mundo» derrotará *con toda probabilidad* a nuestra máquina. Podemos hacer que lo haga para cualquier probabilidad que escojamos (menos la unidad); para cualquier período de tiempo que dejemos a la máquina para realizar su inducción; y para prácticamente cualquier criterio de «derrota» que adoptemos.

Al construir una máquina de inducción, nosotros, los arquitectos de la máquina, debemos decidir qué es lo que constituye su «mundo»; qué cosas han de ser sus sucesos individuales; qué constituye una propiedad o una relación; o, en otras palabras, *qué constituye una repetición.* Y somos nosotros quienes debemos decidir a qué tipo de preguntas queremos que responda la máquina.

Pero eso significa que todos los interrogantes más importantes y difíciles fueron ya resueltos por nosotros cuando construimos el «mundo» y la máquina[11].

14. LA IMPOSIBILIDAD DE UNA LOGICA INDUCTIVA

La máquina inductiva descrita en la sección precedente —con todas las faltas y limitaciones fundamentales que la hacen inaceptable como modelo de inducción— es una versión simplificada y perfeccionada de lo que los lógicos inductivos parecen tener en mente. No cabe duda sobre su relativa simplicidad, si la comparamos con los pesados volúmenes —pesados en todos los aspectos— que se han escrito sobre el asunto. Respecto a que la máquina sea una mejora, me gustaría mencionar un argumento en favor de discutir las teorías inductivas en términos de máquinas, en lugar de hacerlo en términos de lenguajes, un argumento que debería atraer a todos los inductivistas. Es que los gatos y los perros aprenden, decididamente, por experiencia, pero sin un lenguaje simbólico. De ahí que una teoría lógica o lingüística de la inducción será siempre un artificio sumamente poco convincente en comparación con la discusión de posibles máquinas inductivas.

Pero la principal mejora proporcionada por nuestra máquina es que rompe con la concepción tradicional de que la inducción, y en especial alguna que otra forma de la regla inductiva simple, es parte de una lógica inductiva que, a su vez, es una generalización de la lógica deductiva.

He dicho, en una sección anterior, que existe una interpre-

[11] Compárese con los tres últimos párrafos de la sección V de mi artículo «Philosophy of Science: A Personal Report», citado en la nota 9 de la sección anterior, y el Apéndice *X de *L. I. C.*

tación lógica del cálculo de probabilidades que hace de la derivabilidad lógica un caso especial en una relación expresable en términos del cálculo de probabilidades. Por tanto, afirmo la existencia de una lógica probabilística que es una auténtica generalización de la lógica deductiva.

Pero niego la posibilidad de que esta lógica probabilística pueda ser interpretada nunca como una lógica inductiva. Más precisamente, niego la posibilidad de vincular esta lógica probabilística a cualquier forma de la regla inductiva simple.

La existencia del vínculo suele darse por sentada. Esto parece debido a un argumento tácito o explícito en las siguientes líneas.

El razonamiento inductivo es similar en algunos aspectos al razonamiento deductivo, aunque no es tan conclusivo como él. Puede explicarse esta situación suponiendo que un argumento conclusivo (o casi conclusivo) corresponde a la probabilidad 1, un argumento inconclusivo a una probabilidad más baja. Así, el razonamiento inductivo formará parte de la lógica probabilística; y también alguna que otra forma de la regla inductiva simple, que es la forma más elemental y simple de todo el razonamiento inductivo.

Este argumento es completamente equivocado, y la regla inductiva simple no tiene nada que ver con la lógica, ni con la lógica en su forma deductiva habitual, ni con la lógica probabilística a. Y nadie ha desarrollado nunca un argumento que lleve desde la lógica probabilística a ninguna forma de la regla inductiva simple. La mayoría de los escritores sobre el tema han dado por sentado este vínculo; y los pocos que vieron que esto encerraba un problema (especialmente Carnap en su obra *Logical Foundations of Probability*) trataron de salvar la brecha por medio de alguna forma u otra del argumento que he descrito en la sección 12 como «transcendental»: tenía que existir el vínculo porque si no no podríamos aprender nunca por experiencia, según ellos creían.

Es muy fácil probar que la regla inductiva simple tiene que ser independiente de cualquier cosa parecida a la lógica. La prueba de independencia procede, como todas estas pruebas, mediante la construcción de un modelo. Tomemos simplemente, para nuestro mundo, los sucesos del juego «rojo o azul»: en este mundo, la «lógica» continua siendo (objetivamente) válida, como ocurre en todos los mundos lógicamente posibles. Pero la regla inductiva se hace (objetivamente) inválida, en el sentido de que lo más probable será que falle, con una probabilidad tan alta como decidamos.

Vemos ahora por qué nuestra máquina inductiva es tan simple, comparada con casi todos los intentos de construir una

lógica inductiva. La razón es ésta. No hemos tratado de establecer la validez de nuestra máquina por ninguna lógica inductiva ni de establecer un vínculo entre la máquina y la lógica probabilística. No lo hemos hecho porque no puede hacerse. Una lógica inductiva es, precisamente, un intento de establecer lo imposible. Esto es lo que hace tan complicados los libros sobre lógica inductiva y tan flojos sus razonamientos y lo que es responsable de la introducción de interminables y complicados sustitutos del argumento sencillo.

15. LOGICA PROBABILISTICA *VERSUS* LOGICA INDUCTIVA

Hasta aquí he atacado nuestro problema desde el lado de la regla inductiva, estableciendo su falta de validez. Voy a atacarlo ahora desde el otro lado, el de la lógica probabilística.

Se supone que la lógica probabilística es una generalización de la lógica deductiva y sus fórmulas se suponen lógicamente verdaderas, o analíticas o tautológicas, como son las de toda lógica.

Manteniendo presente esto, podemos considerar unas cuantas fórmulas del cálculo de probabilidades que hacen posible esta interpretación.

1) $p\ (a,a) = 1$

Esta fórmula puede interpretarse como «*a* se sigue de *a*».

2) $p\ (a,ab) = 1 = p\ (a,ba) = p\ (a,bac)$

Esta fórmula puede interpretarse como «*a* se sigue de cualquier conjunción de la cual *a* es componente».

3) $p\ (ab) = p\ (a)\ p\ (b)$ si, y sólo si, $p\ (a,b) = p\ (a)$

Podemos interpretar esto como «*a* y *b* son (absolutamente) independientes en el sentido de la lógica probabilística si, y sólo si, la información *b* deja inalterada la probabilidad absoluta de *a*; o, en otras palabras, si *b* no es pertinente a *a*».

4) Si $p\ (ab) \neq p\ (a)\ p\ (b)$ entonces, o bien

i) $p\ (a,b) > p\ (a)$

o bien,

ii) $p\ (a,b) < p\ (a)$ y $p\ (\bar{a},b) > p\ (\bar{a})$

Podemos interpretar esto como «si *a* y *b* no son lógicamente independientes en el sentido de la lógica probabilística, enton-

ces, o bien i) b es favorable a a o apoya a a, ii) b es desfavorable a a o bien debilita a a apoyando a no-a».

Esto muestra que la independencia lógico-probabilística debe interpretarse como neutralidad lógica: «a y b son lógicamente independientes, si, y sólo si, b no favorece ni a a ni $\bar{a}$».

Puede formularse ahora el siguiente teorema, que arroja mucha luz sobre el carácter tautológico de esta lógica[12].

5) Sea a una conjunción de x e y, es decir, $a = xy$, con $p\ (y,x) \neq 1$. (El contenido, tanto de x como de y se supone $\neq 0$.) Sea $p\ (x,b) = p\ (x)\ p\ (b)$, de modo que x y b sean (probabilísticamente) independientes. Sea $b = yz$. Entonces

$$p\ (a,b) = p\ (x,b) = p\ (x) > p\ (a)$$

Es decir, si a es la conjunción de una x que es (probabilísticamente) independiente de b y de una y que se sigue de b, entonces, $p\ (a,b)$, aunque igual a la probabilidad absoluta de x, será, no obstante [siempre que $p\ (y,x) \neq 1$] mayor que la probabilidad absoluta de a[13].

De este modo, 5) nos lleva a la siguiente explicación del carácter tautológico de esta extensión de la lógica deductiva: b favorece a a siempre que a sea equivalente a una conjunción xy

---

[12] Estoy tratando este problema de una forma sencilla, basándome en la posibilidad de analizar contenidos en términos de conjunción. Una discusión más formal tendría que introducir conjunciones de enunciados atómicos relativos (cuya independencia lógica sería entonces el aspecto principal del argumento).

[13] [*(Añadido en enero 1981.) David Miller tiene importantes comentarios sobre esto, dos de los cuales pueden mencionarse aquí.

i) Miller señala que sería interesante ver si alguno de los diversos conceptos de la independencia *lógica* (no-probabilística; tal como la independencia completa; o la independencia máxima: véase Tarski, *Logic, Semantics, Metamathematics*, 1956, p. 36) puede ser relacionado con la independencia probabilística. Y construye efectivamente un vínculo entre ésta última y la independencia máxima: dos enunciados a y b son (según H. M. Sheffer) máximamente independientes si $a \vee b$ es una tautología de tal modo que $p\ (a \vee b) = 1$; y Miller demuestra que, si $p\ (a \vee b) = 1$ (y $p\ (b) \neq 0$), no es posible que b apoye a a: sólo puede debilitarlo o ser (en el mejor de los casos) neutral respecto a a. La prueba de Miller es indirecta: muestra que si $p\ (a,b) > p\ (a)$, entonces, $p\ (a \vee b) < 1$. [De $p\ (a,b) > p\ (a)$ y $p\ (b) \neq 0$ obtenemos $p\ (ab) > p\ (a)\ p\ (b)$ y así, $p\ (a \vee b) < p\ (a) + p\ (b) - p\ (a)\ p\ (b) = p\ (a)\ (1 - p\ (b)) + p\ (b) \leq 1 - p\ (b) + p\ (b)$. Por tanto, $p\ (a \vee b) < 1$.]

ii) Miller usa esta prueba con el fin de establecer el siguiente teorema 6): Si b apoya a a sin entrañarlo, entonces siempre existen dos enunciados x e y tales que $a = xy$ y tales que y resuma todo el contenido de a entrañado por b, mientras que b es, en el mejor de los casos, neutral respecto a x. (Prueba: pongamos $x = a \vee \bar{b}$; $y = a \vee b$. Ahora bien, b no puede apoyar a x puesto que $x \vee b$ es una tautología. A propósito, puede decirse que x e y son máximamente independientes; y pueden llamarse «máximamente independientes con respecto a a».)]

tal que *y* no sea una tautología y se siga de *b*, a condición de que *x* (la parte de *a* no entrañada por *b*) sea independiente de *b* o no muy fuertemente debilitada por *b*: no necesita, en modo alguno, ser apoyada por *b*. Si *b* es independiente de *xy*, por tanto, neutral con respecto a *x*, la probabilidad de *a* dado *b* será igual a la de *x* dado *b*; y, por tanto, a la probabilidad absoluta $p(x) > p(a)$.

Dicho de otro modo, *b* apoya a *a* si entraña parte de su contenido y si es neutral (o no es fuertemente desfavorable) hacia el resto.

Pero si esta interpretación es correcta —y me parece que está claro que lo es—, *entonces la lógica probabilística no puede ser lógica inductiva.* Porque la teoría de la inducción trata de justificar el intento de obtener de las premisas más de lo que hay en ellas: de ir con la condición *a* más allá de la premisa *b*, incluso al precio de una pérdida de certeza. Sin embargo, la probabilidad lógica dice que si el contenido de *a* va más allá de la premisa *b*, entonces sólo la parte de *a* que se sigue de *b* se hace cierta y el resto sigue tan probable o tan improbable como era sin *b*. El aumento en la probabilidad de *a* que puede procurar *b* se debe así, enteramente, al hecho de que parte de *a* está lógicamente entrañada por *b*; no es debido a ningún efecto de *b* sobre la parte de *a* que no se sigue de *b*. [*Por tanto, en la teoría de la probabilidad no hay inferencia probabilística ampliativa. La aparente inferencia ampliativa se debe al hecho de que parte del contenido de *a* que es deducible de *b*.]

Es esta situación la que da a la lógica probabilística su carácter tautológico. Cualquier otra interpretación —digamos que una que hiciera posible que *b* fuera favorable a *a* sin entrañar lógicamente parte de *a*— haría que las fórmulas resultantes fuesen no-analíticas.

La lógica probabilística generaliza así la lógica derivativa al considerar no sólo las conclusiones que son entrañadas por las premisas, sino también conclusiones parciales que son sólo parcialmente entrañadas por las premisas (y también varias otras posibilidades). Pero, al hacerlo, no nos permite proceder desde la premisa conocida —con menos que certeza— a una conclusión desconocida; de hecho, lo que va más allá de la premisa continúa siendo tan probable o tan improbable como era antes.

La inducción, naturalmente, es esencialmente un intento de extender nuestro conocimiento: de proceder de lo conocido a lo desconocido, aumentando al menos la probabilidad de algo que es desconocido. Mi argumento aquí es que, sea lo que fuere lo que pensemos de la inducción no puede identificarse con la lógica probabilística.

16. LA INTEPRETACION INDUCTIVISTA DE LA PROBABILIDAD

¿Qué es, entonces, la «lógica inductiva»? Hay i) una interpretación lógica del cálculo de probabilidades, o «lógica probabilística», que no tiene nada que ver con la inducción; y hay ii) una interpretación inductivista del cálculo de probabilidades que no tiene nada que ver con la lógica —excepto que puede formularse de tal manera que la lógica probabilística sea un caso cero de ella—. Pero *no es una generalización* de la lógica probabilística porque, en la medida en que es inductiva, está en conflicto directo con la interpretación lógica; y si tomamos su caso cero, entonces excluimos cualquier argumento inductivo. Hablar de lógica inductiva es, por tanto, sumamente engañoso. Pero ¿podemos quizá llamar al cálculo de probabilidades, si se interpreta de forma que dé la regla inductiva simple, «*cálculo inductivo*» y a las probabilidades que él calcula, «*probabilidades inductivas*»?

El «cálculo inductivo» es, formalmente considerado, el cálculo de probabilidades más una *regla* que afirma lo siguiente:

Cualquier ejemplo observado de una propiedad P tiende a aumentar (aunque quizá sólo ligeramente) la probabilidad de esa propiedad P. En consecuencia, cualquier ejemplo observado de la propiedad complementaria $\overline{P}$ tiende a disminuir la probabilidad de p (puesto que tiende a aumentar la de $\overline{P}$).

Carnap[14] llama a esto el «teorema de la *pertinencia de los ejemplos*» y lo formula como sigue:

> «... un ejemplo de una propiedad es positivamente pertinente a la predicción de otro ejemplo de la misma propiedad». (Continúa diciendo: «Este parece ser un rasgo básico de todo razonamiento inductivo relativo a la predicción de un suceso futuro».)

La regla significa en la práctica que, por ejemplo, la observación de cosas verdes incrementa la probabilidad de que la próxima cosa sea verde; o que la observación de objetos ingeniosos incrementa la probabilidad de que el próximo objeto observado sea ingenioso. Como lo muestra Carnap, es una forma de lo que yo he llamado «*la regla inductiva simple*».

No quiero tratar aquí la plausibilidad o implausibilidad de esta regla, ni tampoco volver a abrir la cuestión de su validez o falta de validez. Sólo quiero mostrar que, aunque yo no hubiera demostrado, por medio de un contra-ejemplo, que no

[14] R. Carnap, «On the Comparative Concept of Confirmation», *British Journal for the Philosophy of Science* 3, p. 315. (La propia regla se formula como T1 en p. 314, y A4 en p. 316.)

forma parte de la lógica[15], todo el carácter de la regla excluiría su aceptación como parte de la lógica probabilística.

Porque no hay nada ilógico (en el sentido de «autocontradictorio») en la regla opuesta. Me refiero a la famosa *falacia del jugador* de esperar un bloque de «caras» y apostar entonces por «cruces», «porque ha llegado el momento del cambio». Aunque esto es una falacia en lo que concierne a los juegos de azar normales, podemos construir fácilmente una sucesión (digamos que con una frecuencia de los bloques reducida) en la que la falacia del jugador diese buenos resultados.

Pero la formulación de la regla inductiva que acabo de dar muestra su carácter no lógico, no sólo por estar en conflicto con la falacia del jugador o con nuestro juego de «rojo o azul»; está en conflicto con cualquier sucesión aleatoria normal, ya que, interpretada objetivamente, postula un efecto secundario que en una sucesión aleatoria normal no ocurre. La regla choca, pues, con la estimación

$$p\,(a,b) = 1/2$$

por ejemplo, donde *a* es sacar cara con una moneda homogénea, como nos dice *b*.

Este hecho queda oscurecido por lo que he llamado más arriba la «ley de la utilidad decreciente» que asegura que la probabilidad, calculada por medio de la regla inductiva simple, se estabiliza muy cerca del valor «verdadero» u «objetivo de la probabilidad» de una sucesión aleatoria, como hemos visto, llevando así, finalmente, al «éxito». Pero incluso después de haberse estabilizado de este modo, la probabilidad inducida puede oscilar un poco si aparece un bloque inusualmente largo (no necesita ser un bloque puro), incluso aunque este bloque llegue algo retrasado desde un punto de vista objetivo.

El hecho de que todas las formas de la regla inductiva simple estén en conflicto con la probabilidad objetiva, incluso para el juego corriente de tirar a cara o cruz, merece ser bien destacado. El conflicto puede ser secundario al empezar y acabará por desaparecer totalmente si la sucesión de jugadas se hace infinitamente larga. Pero esto no cambia el hecho de que si hay implicadas sumas elevadas, cualquier jugador que utilice una forma de la regla inductiva simple como su «guía para la vida»

[15] Aparte del juego «rojo o azul» he construido también, en *L. I. C.*, Apéndice *VII, nota 9, un ejemplo de propiedad *B* que contradice el «*teorema de la pertinencia de los ejemplos*». Hay que admitir que *B* sólo puede construirse en un lenguaje más rico que el de Carnap; pero una regla que, si se extiende a un lenguaje más rico, lleva una contradicción, no puede ser considerada un principio lógico ni analítico; especialmente si no hay argumento en su favor (excepto el argumento «trascendental», *cf.* nota 3 al Apéndice *VII).

o, en otras palabras, como su *sistema de juego*, deberá esperar grandes pérdidas. Porque la probabilidad calculada según la regla inductiva simple se desviará una y otra vez de la probabilidad objetiva (que ponemos en 1/2, ya que estamos hablando de tiradas con una moneda). Siempre que ocurra esta desviación, el jugador inductivo debe estar preparado a aceptar probabilidades que, desde el punto de vista objetivo, son desfavorables. Pero, aunque esto no tiene que llevar necesariamente a pérdidas, la pérdida (objetivamente) esperada o probable puede hacerse tan grande como gustemos: una clara indicación de que las probabilidades inductivas no pueden ser tautológicas.

Nuestra discusión establece, creo, dos cosas: 1) la coherencia del cálculo inductivo (interpretado), es decir, de la probabilidad inductiva[16]; 2) su carácter no lógico. Realmente, uno no debería argüir ya sobre este último punto, 216 años despues del *Treatise* de Hume; pero los pasajes pertinentes no han recibido, evidentemente, la atención que merecen. Hume arguyó que no podía haber argumento demostrativo (o lógico) válido que nos permitiera mostrar «*que aquellos casos de los que no tenemos experiencia se asemejan a aquellos de los cuales tenemos experiencia*». Por consiguiente, «*incluso tras la observación de la frecuente o constante conjunción de objetos, no tenemos razón alguna para hacer ninguna inferencia que concierna a ningún objeto fuera de aquellos de los que hemos tenido experiencia*». Pues, «si dijera que tenemos experiencia» —experiencia que nos enseña que los objetos en conjunción constante con otros objetos, continúan estando en la misma conjunción— entonces, dice Hume, «yo reiteraría mi pregunta, *por qué, de esta experiencia, sacamos una conclusión más allá de aquellos casos pasados de los cuales hemos tenido experiencia*»[17]. Dicho de otro

[16] [*El cálculo de la probabilidad inductiva es coherente sólo si se aplica una sucesión estabilizada (suficientemente larga) de ocurrencias objetivamente independientes; pero es incoherente en la medida en que afirma la pertinencia (y, por tanto, lo contrario de la independencia) de esas ocurrencias, pero si la sucesión se estabiliza, *esta* incoherencia teórica no afecta a las aplicaciones. (Naturalmente, el juego de «rojo o azul» muestra que no debe aplicarse este cálculo a sucesiones objetivamente dependientes.) Véase también David Miller, «A Paradox of Information», *British Journal for the Philosophy of Science* 17, 1966, pp. 59-61.]

[17] Las cursivas son de Hume. *Cf.* Hume, *Treatise on Human Nature*, 1739-40, libro I, parte III, secciones VI y XII. (Edición de Selby-Bigge, pp. 89, 139 y 91.) Para Hume era muy claro que el recurso a la probabilidad no alteraría la situación. Así, escribe en su *Abstract*: «Es evidente que *Adán*, con toda su ciencia, nunca hubiera sido capaz de *demostrar* que el curso de la naturaleza debía continuar siendo *uniformemente* el mismo... No, iré más lejos, afirmaré que no podía probar con ningún argumento *probable* que el futuro debía conformarse al pasado. Todos los argumentos probables se basan en la proposición de que existe esta conformidad entre el futuro y el pasado y, por tanto,

modo, el intento de justificar la práctica de la lógica inductiva es insostenible y el intento de justificarla por medio de una apelación a la experiencia tiene que llevar a una *retrogresión infinita.*

Por tanto, la regla de Carnap que establece la dependencia mutua, o pertinencia, de dos casos [ejemplos] cualesquiera de la misma propiedad es un principio inductivo no-analítico; y si se sostiene (y si hay que evitar la retrogresión infinita de Hume), sólo puede sostenerse como proposición sintética, es decir, como válida *a priori.*

¿Cómo podemos, entonces, argüir en su favor? Con el argumento transcendental, desde luego. Esto está ahora muy de moda entre todo tipo de lógicos inductivos que desdeñan a Kant por *apriorista* y transcendentalista. (Yo soy un gran admirador de Kant, pero creo que el uso positivo del argumento transcendental en apoyo de principios que él creía válidos *a priori* fue su defecto más grave. Sin embargo, había muchas circunstancias atenuantes en aquel tiempo, tan anterior a Einstein)[18]. Puede encontrarse el argumento referido a diversas teorías probabilísticas de la inducción, en Russell, Jeffreys y Reichenbach. Carnap también lo usa, pero con una cierta vacilación que, en mi opinión, dice mucho en favor de este autor. Porque escribe sobre su elección de la función m^* (que equivale a la elección de una cierta forma especial de la regla inductiva simple):

> Las consideraciones precedentes muestran que puede presentarse en favor de m^* el siguiente argumento que, admitámoslo, no es un argumento fuerte. De las dos funciones m que son más sencillas y que se nos aparecen como las más naturales, m^* es la única que no es enteramente inadecuada[19].

Escoger la otra función m, que es «enteramente inadecuada» equivaldría a escoger lo que ha llamado *probabilidad lógica.* Su «falta de adecuación» consiste, naturalmente, en el hecho de que no implica ninguna forma de la regla inductiva simple. Pero, desde mi punto de vista, esto no puede ser de otro modo, *puesto que la lógica probabilística no es inductiva; la probabilidad inductiva no es lógica; y puesto que no puede existir una lógica inductiva.*

Pero, aunque no puede existir y no existe una lógica inductiva, sí que existen los lógicos inductivos. De hecho, han sido

---

nunca pueden probarla». *Cf. An Abstract of a book lately published entitled a Treatise on Human Nature,* 1740, editado por J. M. Keynes y P. Sraffa, 1938, p. 15. (Las cursivas son de Hume.)

[18] [Véase *El desarrollo del conocimiento científico (C. & R.),* Introducción y capítulo 8. *Ed.*]

[19] *Logical Foundations of Probability*, p. 565.

bastante abundantes desde Bacon (si no desde Aristóteles). Y no se les refuta fácilmente. Esto ocurre porque las probabilidades inductivas —es decir, los cálculos inductivos aplicados— pueden ser coherentes, incluso aunque no sean tautológicos. Sin embargo, *sólo* serán coherentes mientras se distinga claramente la probabilidad inductiva de la probabilidad lógica y de la probabilidad objetiva. Porque los valores probabilísticos de la probabilidad inductiva son incompatibles, en general, tanto con la probabilidad lógica como con la probabilidad objetiva, incluso de un juego simétrico e independiente tal como el de tirar monedas, como acabamos de ver. Apenas necesito añadir que algunos o todos esos diferentes tipos de probabilidades son, por regla general, erróneamente identificados en los argumentos transcendentales de los lógicos inductivos[20].

17. LA SUPERFLUIDAD DE LAS TEORIAS

El método de hipótesis que se defiende en *L. I. C.* es muy viejo. Lo que puede ser nuevo es la actitud hacia las contrastaciones predictivas. En ocasiones se ha considerado que éstas eran verificaciones o incluso pruebas de una teoría. Más tarde se ha considerado que eran «simplemente» capaces de hacer probable la teoría. Yo las considero tentativas de falsación; y sólo si, a pesar de ser tentativas sinceras e ingeniosas, no logran refutar la teoría, las considero capaces de corroborarla. Al corroborar las teorías, pueden hacerlas parecer «probables», pero en ninguno de los sentidos de la palabra «probable» que satisfacen al cálculo de probabilidades: por esta razón, prefiero hablar de «corroboración» (o «confirmación» o «aceptabilidad»).

La importancia del método de hipótesis o del método hipotético-deductivo es ahora generalmente reconocida e incluso los inductivistas admiten plenamente que una teoría de la inducción que no diese a este método toda la importancia que merece, sería disparatada. Algunos de los principales inductivistas, como Jeffreys o Carnap, se han esforzado por poner de relieve este punto. Jeffreys pone la cuestión de las teorías universales y de su simplicidad en el centro de su teoría

[20] Quiero dejar completamente claro, en esta ocasión, que mi «grado de corroboración» (o confirmación o aceptabilidad) puede ser definido en términos de probabilidad *lógica*, que resulta ser, de ese modo, perfectamente adecuada para ese propósito. Pero no es idéntico, naturalmente, a la probabilidad lógica ni a ninguna otra función que satisfaga el cálculo de probabilidades.

de la inducción. Y Carnap resalta, en la parte introductoria de su libro, que con «inducción» no quiere decir un método que recoja datos y los generalice, sino un método para evaluar «el grado de confirmación» que obtiene una teoría inventada alibremente, por medio de contrastaciones empíricas.

> Si, por ejemplo —escribe Carnap— tenemos un informe sobre los resultados de una observación y queremos encontrar una hipótesis que esté bien confirmada y proporcione una buena explicación de los sucesos observados, no existe un conjunto de reglas fijas que lleve automáticamente a la mejor hipótesis, ni siquiera a una buena. *Es cuestión de ingenio y suerte, el que el científico dé con una buena hipótesis; y si encuentra una*, nunca puede estar seguro de que no pueda haber otra hipótesis que encaje aún mejor con los hechos observados, antes incluso de realizar nuevas informaciones. Este punto, la imposibilidad de un procedimiento inductivo automático, ha sido especialmente resaltado, entre otros, por Karl Popper... que tambien cita un enunciado de Einstein...[21]

Un poco más adelante, escribe:

> Dada: una cláusula *e* como evidencia: buscada: una hipótesis *h* que que esté muy confirmada... No existe procedimiento efectivo para resolver estos problemas: éste es el punto que han resaltado Einstein y Popper, como he mencionado más arriba.

El problema que debería intentar resolver la lógica inductiva es, según el programa de Carnap, no el descubrimiento de hipótesis, sino la determinación de su grado de confirmación:

> Dadas: dos cláusulas *e* y *h*. Buscado: el valor de ...el grado de confirmación de *h* sobre la base de la evidencia *e*.

Ahora bien, está claro que este programa toma debidamente en cuenta el método de las hipótesis; y yo estoy totalmente de acuerdo con todo lo que ha dicho Carnap en los pasajes citados (excepto, naturalmente, que no puedo aceptar la interpretación del «grado de confirmación» como probabilidad en el sentido de cálculo de probabilidades y menos aún como una «probabilidad inductiva»).

Pero este bien defendido programa choca de la forma más asombrosa con su ejecución. Primero, a lo largo de su libro, Carnap introduce un método que nos permite «descubrir» o, más bien, calcular, la hipótesis *mejor* confirmada, es decir, la hipótesis más probable basada en cualquier evidencia dada. Pero esto singnifica, naturalmente, que ya no somos libres de inventar una hipótesis; o, más bien, nunca tuvo ningún sentido que nos devanásemos los sesos tratando de encontrar una

[21] *Logical Foundations of Probability*, pp. 192 s. (Los subrayados son míos.) Los dos pasajes siguientes son de pp. 194 s. y 196.

buena hipótesis, puesto que, simplemente, podíamos haber calculado la mejor hipótesis.

Así, el método de hipótesis se convierte en totalmente superfluo, en directa contradicción con el programa de Carnap. Y, sin rechazar explícitamente su programa original, el propio Carnap admite su superfluidad. Al final de su libro, en una sección titulada «¿Hacen falta leyes para hacer predicciones?» (pp. 574 ss.), escribe, después de dar un resumen de sus resultados y después de mostrar que las «leyes» o teorías universales no le hacen realmente falta al científico (Mr. X):

> Vemos así que X no necesita, en absoluto, tomar el tortuoso camino a través de la ley *l*, como suele creerse; en su lugar, puede ir desde su conocimiento observacional... directamente a la predicción singular...

Continúa:

> La forma acostumbrada de pensar en la vida cotidiana toma a menudo este mismo atajo que está justificado por la lógica inductiva [22].

Y concluye:

> Vemos que el uso de leyes no es indispensable para hacer predicciones. No obstante, es conveniente, desde luego, enunciar leyes universales en los libros de física, biología, psicología, etc.

Como, según esta filosofía de la ciencia, la ciencia pura no tiene otra tarea que la de predecir, esto significa que las leyes científicas no tienen función alguna; son completamente superfluas. De hecho, el argumento de Carnap muestra que son desvíos innecesarios y, por tanto, deberían ser abandonados aplicándoles o bien la navaja de Ockham o el principio de economía de Mach (que resulta que fue utilizado por Mach originariamente para mostrar la utilidad de las leyes universales). Pero todo esto significa la superfluidad del método hipotético-deductivo. Ya no hace falta «ingenio y suerte» para «dar con una hipótesis adecuada»: el tipo de hipótesis adecuada a la que se refería era una teoría universal y las teorías universales ya no son «necesarias»; ya no son «indispensables». Brevemente, la creencia de que el método de la ciencia era el método de las hipótesis era toda ella un error.

Me siento inclinado a aplicar a esta conclusión el argumento transcendental, aunque sólo sea en su versión negativa. Una teoría de la ciencia que rechaza el método hipotético-deductivo diciéndonos que las teorías científicas son superfluas

[22] No necesito decir hasta qué punto estoy en desacuerdo con esta aserción, pero puede merecer la pena referirse aquí a una afirmación similar de J. S. Mill, *Logic*, libro II, cap. III, 3: «Toda inferencia es de particulares a particulares».

es «enteramente inadecuada», precisamente en el mismo sentido en que lo es una teoría que nos dice que no podemos aprender por experiencia.

18. CARECE DE SENTIDO CONTRASTAR UNA TEORIA

La conclusión que he rechazado en la sección anterior había sido explícitamente formulada por Carnap, pero no es, en absoluto, peculiar de su teoría [23]. De uno u otro modo, las diversas teorías probabilísticas de la inducción establecen, todas ellas, demasiado y en la mayoría de los casos, más de lo que se proponen; no se reducen a dar una estimación del grado en el cual una teoría ha sido contrastada, sino que dictan, en cada caso, qué teoría debemos aceptar como la mejor teoría, es decir, como la más probable. Como consecuencia, la cuestión de *contrastar* una teoría no se plantea ya. Se hace inútil. La inducción no resulta ser un método de contrastar una conjetura o de evaluar el resultado de una contrastación; en cambio, se convierte en un método para calcular la mejor teoría por medio de una inferencia inductiva [24].

Esto se debe a la aplicación de la regla simple de inducción en cualquiera de sus diversas formas (véase la sección 5, nota 4, *supra*).

La regla simple de la inducción nos dice cuál es la probabilidad del próximo caso. Si con «*a*» nos referimos al enunciado de que el próximo caso o ejemplo tiene la propiedad *A* y con «*b*», a nuestra experiencia pasada que implica $m/r = r$ casos de *A*, entonces la regla simple nos dice que tenemos, aproximadamente

$$p\,(a,b) = m/n = r$$

y la aproximación se hace mejor al aumentar n.

En la medida en que esta fórmula interesa a un científico, su punto principal es que nos puede permitir derivar una *hipótesis sobre frecuencias*. Esto puede hacerse, si podemos suponer

[23] Russell se acerca a una formulación explícita en su *Human Knowledge*, 1948, p. 435 (parte V, VII, E).

[24] [*(Añadido en 1980.) Según la explicación de Carnap (tal como la desarrolla en *Logical Foundations of Probability*), la lógica probabilística, aunque es una generalización de la lógica deductiva, no consiste «en hacer inferencias sino, más bien, en atribuir probabilidades» (véase sus observaciones en I. Lakatos, ed., *The Problem of Inductive Logic*, 1968, p. 311). Sin embargo, la razón para atribuir probabilidades es, evidentemente, poder escoger de una clase de hipótesis posibles, la hipótesis a la que se hayan atribuido probabilidades más elevadas. Pero este proceso de elegir puede describirse, desde luego, como el proceso de hacer una inferencia inductiva.]

independencia, por medio del teorema de Bernoulli. Sin este teorema habría que postular que «*la frecuencia más probable*» de ejemplos con la propiedad *A* en cualquier clase de ejemplos es igual a la probabilidad que hemos calculado por medio de la regla inductiva, es decir, igual a *r*.

Podemos llamar a *r*, entonces, «la mejor estimación de la frecuencia de *A* en cualquier clase futura».

Si queremos distinguir por medio de nuestro formalismo la probabilidad inductiva

$$p\,(a,b) = r$$

de la mejor estimación de la frecuencia correspondiente (que es igual a ella), entonces podemos indicar la *probabilidad inductiva*, calculada por medio de la regla inductiva, con «p_i» y podemos escribir

$$p_i\,(a,b) = r$$

para «la propabilidad inductiva (dado *b*) de que el próximo ejemplo tenga la propiedad *A* es igual a *r*».

En contraposición, podemos escribir «p_s» para la mejor estimación de la *probabilidad estadística*, o frecuencia relativa. Para simplificar nuestra simbología, podemos acordar leer la letra «*a*» en una expresión de p_s de modo ligeramente diferente a la letra «*a*» en una expresión de p_i. Es decir, escribiremos

$$p_s\,(a,b) = r$$

con el significado: «la mejor estimación (dado *b*) de la probabilidad estadística o frecuencia relativa, de ejemplos con la propiedad *A* en cualquier clase[25] futura de ejemplos, es igual a *r*».

Ahora bien, si el teorema de Bernoulli es aplicable, obtenemos

1) $$p_s\,(a,b) = p_i\,(a,b);$$

y si no podemos probar esto (porque no sabemos lo suficiente sobre la independencia de las condiciones descritas por *b*, por ejemplo, si *b* exige que se mezclen las bolas en una urna) todavía podemos introducir la fórmula 1) como *postulado* o quizá por medio de una definición. Puede expresarse, más brevemente, como

2) $$p_s = p_i$$

[25] No discutiré el problema que plantea la frase «*cualquier* clase futura» (aunque esta frase lleva a consecuencias paradójicas) porque otras críticas que quiero hacer me parecen más fundamentales.

Esta fórmula es la base de todas las estimaciones inductivas de frecuencias, al menos de las que yo tengo conocimiento[26]. (Naturalmente, si $p_s = p_i$ se introduce como postulado, equivale a postular que *b* exige independencia.)

Ahora todas nuestras hipótesis objetivas pueden interpretarse como estimaciones de frecuencias. (No es que yo recomiende ahora precisamente esta interpretación. Pero esto no importa aquí; además obtendríamos precisamente el mismo resultado para «p_p» en vez de «p_s», es decir, para la mejor estimación de la propensión.) Así, puede decirse que 1) y 2) *explican las hipótesis de la probabilidad objetiva en términos de probabilidad subjetiva*. Hacerlo era, después de todo, la principal tarea de la teoría subjetiva; y 1) expresa el (supuesto) hecho de que su tarea ha sido realizada. De este modo, una fórmula como 1) está esencialmente implicada en todas las teorías subjetivas, aunque rara vez está enunciada de forma explícita (como lo está en el libro de Carnap)[27].

Los lógicos inductivos creen en una lógica inductiva, y, por consiguiente, interpretan la probabilidad inductiva como probabilidad *lógica*. Esto significa que, según su concepción, toda *fórmula verdadera* de la forma

$$p_i\,(a,b) = r$$

tiene que se tautológica o analítica y toda fórmula *falsa* de esta forma, auto-contradictoria.

Ahora podemos derivar inmediatamente de 1)

3) $\qquad p_s\,(a,b) = r$ si, y sólo si, $p_i\,(a,b) = r$

Esta fórmula tiene que ser analítica si 1) es analítica; y encontramos así que: si p_i es no sólo inductiva, sino también una probabilidad lógica, entonces toda fórmula de la forma

4) $\qquad p_s\,(a,b) = r$

debe ser o tautológica (analítica) o auto-contradictoria; si es verdadera, tiene que ser tautológica.

---

[26] La fórmula T106 — 1, + *c*, en p. 551 de *Logical Foundations of Probability* de Carnap, significa exactamente lo mismo que nuestras 1) ó 2) si se traduce a nuestro formalismo. En esta fórmula, Carnap no hace uso de sus expresiones «probabilidad$_1$» (o «p_1») que corresponde a nuestro «p_i»; y «probabilidad$_2$» (o «p_2») que corresponde a nuestro «p_s», si consideramos sólo la «*mejor estimación*». La fórmula T106 — 1, + *c* significa, en efecto, algo parecido a «$p_1 = p_2$» (más excactamente, «la mejor p_2 es igual a p_1») como lo muestra mi traducción 2). La definición de Carnap que establece su ecuación es + D 100 — 1 en p. 525; *cf.* 3) en p. 169.

[27] Véase la nota anterior. La derivación de la fórmula del libro de Carnap es sumamente compleja.

Pero esto implica que *todas nuestras «mejores» hipótesis de probabilidades objetivas son tautológicas.*

Esta conclusión es claramente absurda[28]. Porque destruye la posibilidad de que haya auténticas hipótesis: las estimaciones como éstas no pueden *contrastarse* por medio de experimentos.

Podemos tomar esto como una refutación de la concepción de que la probabilidad inductiva es probabilidad lógica o, dicho con otras palabras, que todas las fórmulas verdaderas del cálculo inductivo son tautológicas.

Contra esta crítica puede objetarse que de 4) podemos obtener un enunciado auténticamente empírico si le aplicamos la «regla de absolución» que se trató en la sección 11. Esta objeción contiene la sugerencia tácita[29] de que la «regla de absolución» debe extenderse a las *estimaciones estadísticas.* Dicha extensión, que constituiría un procedimiento muy dudoso, significaría que, al ser *b* nuestro conocimiento *presente* total, obtendríamos de *b* y $p_s(a,b) = r$, por medio de la «regla de absolución», lo siguiente:

5) «En el momento actual, la mejor estimación de la frecuencia (dentro de cualquier clase de ejemplos) de ejemplos con la propiedad *A* es que esta frecuencia es igual a *r*.»

Discutamos brevemente esta sugerencia y sus consecuencias.

i) El enunciado 5) es extramadamente raro desde el punto de vista de la teoría frecuencial, porque debido a la aplicación de la regla de absolución, *no se enuncian las condiciones para el resultado A.* (Dicho de otro modo, se admite cualquier clase de referencia, lo que, o lleva a contradicciones o hace inaplicable la teoría frecuencial. Esta dificultad es fundamental y no se puede eliminar considerando a *a* como enunciado condicional y a *A* como propiedad condicional.)

ii) Incluso aunque dejemos a un lado todos estos problemas, la siguiente situación permanece:

Admitamos que 5) se hace empírica con la sugerencia que hemos discutido: se ha obtenido a partir de 4) aplicándole la regla de absolución junto con un informe empírico —el informe de que *b* es nuestro conocimiento presente total—. Pero esto

---

[28] Es exactamente análogo a la situación en otras epistemologías subjetivistas, por ejemplo, en la que se encuentra en Carnap, *Der Logische Aufbau der Welt*, como he mostrado en mi artículo «The Demarcation Between Science and Metaphysics» en el volumen sobre Carnap en la *Library of Living Philosophers*, ed. por P. A. Schilpp; véase ahí la sección 2, texto correspondiente a la nota 27; véase también *El desarrollo del conocimiento científico (C. & R.)*, cap. 11; el lenguaje de *Logische Aufbau* contiene sólo enunciados que son o tautológicos o contradictorios.

[29] He encontrado esta sugerencia tácita en el artículo de Carnap titulado «Probability as a Guide in Life», *Journal of Philosophy* 44, 1947, p. 144 (final del segundo párrafo).

significa que 5) es *entrañado lógicamente por este informe empírico.* Por tanto, no es una hipótesis ni una atrevida invención ni una conjetura ni una suposición que pueda someterse a contrastaciones. Por el contrario, nos es *impuesta* por nuestra experiencia pasada; puede ser *inferida* de nuestra experiencia pasada.

Pero, si esto es así, si 5) es entrañada lógicamente, con certeza, por nuestro conocimiento pasado, entonces no puede ser más que parte de ese conocimiento. Por tanto, 5) *no habla, en absoluto, sobre casos futuros,* a pesar de su formulación verbal. Es, simplemente, la consecuencia de atribuir, por una decisión totalmente arbitraria, el nombre de «mejor estimación del futuro» a un informe sumario sobre el pasado.

Todos éstos son los frutos de intentar escapar de las desagradables consecuencias de la concepción de que 4) es analítica. Si abandonamos esta concepción, llegamos a una teoría que es, esencialmente, la de Reichenbach. Según Reichenbach, debemos aceptar la regla inductiva, o la fórmula 1), basándonos en el argumento transcendental. Todo el procedimiento de llegar a una estimación $p_s\,(a,b) = m/n = r$ no es, como hace notar él, un procedimiento tautológico; por el contrario, él lo llama «*una apuesta*». Con ello quiere resaltar que corremos un grave riesgo al adoptar $p_s\,(a,b) = m/n = r$ como nuestra mejor estimación. Sin embargo, arguye transcendentalmente, *tenemos* que adoptarla porque es la única vía razonable que tenemos abierta, la *mejor* vía. Toda otra estimación sería peor, y arbitraria. Esta es una concepción que hace de

$$p_s\,(a,b) = r$$

una auténtica hipótesis, con la incertidumbre de una auténtica hipótesis; y escapa así a nuestra crítica de que es analítica y, *por tanto,* incontrastable.

Con todo, aunque «las apuestas» de Reichenbach no son, ciertamente, analíticas, no son más contrastables que si fuesen analíticas.

Porque, ¿en qué consistiría una contrastación? Tendría que producir otro ejemplo, que tendría, o no, la propiedad A. Si tiene la propiedad A, entonces obtenemos un r nuevamente determinado tal que $r = (m + 1)/(n + 1)$. Si no tiene la propiedad A, entonces r es igual a $m/(n + 1)$. En ambos casos tenemos que volver a adoptar el valor determinado por 1). Ninguna contrastación nos llevará nunca a ninguna otra parte. Una nueva contrastación sólo cambia b y lleva automáticamente a ese valor de $p_s\,(a,b)$ que corresponde al nuevo b.

Así, una vez más, se abandona el método hipotético. No podemos conjeturar libremente r y contrastar luego nuestra

conjetura; en cambio, el valor de r está determinado únicamente por experiencias pasadas. Las teorías no son hipótesis libremente inventadas: son inferidas por medio de una regla inductiva. Por tanto, la crítica transcendental con que concluía la sección precedente, vuelve a ser aplicable aquí. Tenemos que rechazar un análisis del método científico que hace absurdo el método hipotético.

19. RESUMEN DE ESTA CRITICA

En mi crítica de la teoría subjetiva de la probabilidad, he tratado de ignorar los defectos menores y concentrarme en sus ideas principales: su distinción entre *conocimiento completo* (si podemos derivar a o $\bar{a}$ de la evidencia b) y el *conocimiento parcial o conocimiento probable* (si la evidencia b es insuficiente para obtener un conocimiento completo); su interpretación de la probabilidad como una generalización de la lógica deductiva; y su otra interpretación, muy diferente, de la *inducción* como generalización de la deducción (y, por tanto, de la lógica); su consiguiente identificación de la probabilidad inductiva y la probabilidad lógica; su tentativa de obtener un principio de inducción en forma de un principio de la probabilidad inductiva (tal como la regla inductiva simple), e interpretarlo como un principio de la probabilidad lógica; y finalmente, su conclusión, tal vez no buscada, de que las hipótesis probabilísticas predictivas están totalmente determinadas por nuestras observaciones pasadas, de acuerdo con la regla inductiva simple (siempre que queramos, como sin duda queremos, obtener las mejores hipótesis o «estimaciones»).

Este es el programa subjetivista. Con una excepción, cada una de esas ideas es equivocada. Estoy dispuesto a conceder que hay una *interpretación lógica*, sostenible, del cálculo de probabilidades que lo considera como una generalización de la lógica deductiva. Pero esta interpretación lógica no ofrece apoyo para la inducción; y todas las demás ideas pertenecientes al programa subjetivista llevan a contradicciones y a paradojas.

He omitido aquí muchos puntos de una crítica detallada que he expuesto en otros lugares[30], y una serie de paradojas que tienen, en comparación, un interés limitado.

[30] Véanse mi artículo «The Demarcation Between Science and Metaphysics», citado *supra*, nota 3 a la sección 19; mis tres artículos en *British Journal for the Philosophy of Science* 5, 1954, pp. 143 ss.; 6, 1955, pp. 157 ss. y 7, 1956, pp. 249 ss.; y los Apéndices *VII a *IX de *L. I. C.* Véase también «On Carnap's Versión of Laplac's Rule of Succession», *Mind* 71, 1962, pp. 69-73, y «The Mysteries of Udolpho: A Reply to Professor Jeffrey and Bar-Hillel», *Mind* 76, 1967, pp. 103-110.

He emprendido esta tarea no porque quiera combatir ciertas ideas filosóficas, sino principalmente a causa del sorprendente papel que ha jugado la teoría subjetiva dentro de la propia física: en física y en la filosofía de la física, el modo de discutir problemas probabilísticos consiste, en gran medida, en un cambio continuo de problemas objetivos a argumentos subjetivos y luego otra vez a soluciones objetivas. Sin este cambio continuo de una interpretación a la otra, la debilidad de la teoría subjetiva hubiera sido descubierta hace mucho tiempo. Esta oscilación es tan frecuente entre físicos (incluido Einstein) como entre filósofos.

El papel que ha jugado la teoría subjetiva en la física puede explicarse, en gran medida, por el hecho de que continúan teniendo influencia las concepciones deterministas o restos de ellas, incluso entre los defensores del indeterminismo: el determinismo lleva a la interpretación subjetiva de la probabilidad en física, simplemente porque no deja ningún espacio para una interpretación objetiva de la probabilidad[31].

La debilidad de la teoría subjetiva de la probabilidad la comparte toda epistemología subjetivista (sensualista, fenomenalista, solipsista, etc.). Con epistemología subjetivista me refiero al intento de responder a la pregunta: «¿cómo lo sabe?»[32] en el sentido de[33] «¿cuál es la base de su aserción? ¿qué observaciones le llevaron [a ella]...?». Las preguntas subjetivistas e inductivistas como esas piden las respuestas subjetivistas e inductivistas usuales. Mi propia respuesta sería: «*No* sé: mi aserción no era más que una suposición. No importa qué observaciones puedan haberme llevado a ella. En cambio, puede usted ayudarme criticando mi aserción y utilizando su ingenio para idear algunas contrastaciones experimentales con las que refutar mi aserción si es errónea, como puede muy bien serlo»[34].

Las meras suposiciones, sin embargo, siguen sin estar de moda. Las teorías científicas se suponen «basadas» directamente en la experiencia pasada y, en el mejor de los casos, se admite que la inducción implica una apuesta. No veo por qué es mejor una apuesta que una suposición; sobre todo si recor-

---

[31] [Véanse los tomos II y III del *Post Scriptum*, en los que los problemas de determinismo y física se tratan con detalle. *Ed.*]

[32] También Jeffreys hace esta pregunta en el sentido enunciado aquí, en *Theory of Probability*, 1.ª edición, p. 34; 2.ª edición, p. 33.

[33] Las dos preguntas-guías siguientes las hace Carnap en su *Logical Foundations*, p. 189; véase también la sección 4 *supra*.

[34] Véase también la sección 27 de *L. I. C.*, especialmente notas 2 y *1, y el texto correspondiente. [Véase también «Sobre las fuentes del conocimiento y de la ignorancia», en *El desarrollo del conocimiento científico (C. & R.)*.]

damos que la teoría subjetiva-inductiva empieza a parecerse un poco a una apuesta a ciegas.

Pero el tema general del subjetivismo nos apartaría de nuestra tarea real en esta Parte: el análisis de la probabilidad.

**Addendum (enero de 1981). Un breve resumen de la crítica de la inducción probabilística*

I

Lo que persuade a tanta gente a aceptar la inducción probabilística son dos cosas:

A) La idea intuitiva *válida* pero *mal interpretada* de que *no puede deberse solamente a un accidente improbable* el que una hipótesis tenga éxito una y otra vez cuando se la somete a contrastaciones en circunstancias diferentes y, especialmente, si tiene éxito haciendo predicciones previamente inesperadas. (*Si no se debe a un accidente improbable*, nos inclinamos a razonar, de modo no válido, entonces tiene que deberse a *la alta probabilidad de la hipótesis*. Véase *infra*.)

B) El hecho indudable de que, según el cálculo de probabilidades, la probabilidad de un enunciado aumenta (a menos que fuese cero al principio) con la acumulación de evidencia en su favor, sobre todo con la acumulación de predicciones de éxito. En términos del cálculo de probabilidades: sea h una hipótesis cuya probabilidad inicial (o probabilidad previa) $p(h)$ puede ser tan pequeña como queramos, siempre que sea distinta de cero. Sea e cierta evidencia en su favor. Entonces, no sólo tenemos

$$p\,(h,e) > p\,(h)$$

sino también

$$p\,(h,e_2) > p\,(h,e_1)$$

si e_2 contiene evidencia favorable no contenida aún en e_1. De esto modo, la probabilidad de h continúa aumentando con la acumulación de evidencia favorable.

Este hecho, innegable y persuasivo, que se sigue del cálculo de probabilidades, ha convencido a muchas personas para que pensaran que la probabilidad $p\,(h,e)$ de cualquier hipótesis h tiene que tender a 1 con la acumulación de evidencia favorable e.

Los argumentos A) y B) le han parecido contundentes a mucha gente, incluidos muchos filósofos de la ciencia.

Mostraré primero que B), aunque es válido, induce a error. B) se apoya en la comparación del cambio de la probabilidad de *una* hipótesis h cuando cambia la evidencia. Las cosas adquieren un aspecto totalmente distinto si comparamos *dos o más* hipótesis con una evidencia constante e favorable a todas ellas. Entonces encontramos, inmediatamente, que la influencia ejercida por la evidencia sobre la probabilidad no tiene nada que ver, en absoluto, con la inducción.

Esto lo mostraré en la sección II.

II

Sea $h_1, h_2, \ldots h_n$ una sucesión de hipótesis rivales e incompatibles tomadas de dos en dos. Sea e la evidencia pertinente (las predicciones de éxito) derivable de cada una de las n hipótesis, de modo que tengamos que $p(h_i,e) = p(h_i)$ y $0 \neq p(e) \neq 1$. Obtenemos entonces, inmediatamente, a partir del teorema general de la multiplicación del cálculo de probabilidades [véase *L. I. C.*, p. 267 1)], como aquí en la sección I, *supra*,

1) $$p(h_i,e) > p(h_i)$$

es decir, la evidencia favorable e incrementa la probabilidad.

Pero también obtenemos el teorema trivial, aunque fundamental, 2)

2) $p(h_i,e) < p(h_j,e)$ si, y sólo si, $p(h_i) < p(h_j)$

El teorema 2) es aniquilador. Muestra que la evidencia favorable e, aunque eleva la probabilidad según 1), no obstante, en contra de las primeras apariencias, deja todo exactamente como estaba. Nunca puede favorecer a h_i en detrimento de h_j. Por el contrario, el orden que atribuimos a nuestras hipótesis antes de la evidencia, permanece. Es inconmovible por ninguna evidencia favorable. La evidencia no puede influenciarlo.

¿Por qué es aniquilador esto? Porque h_1 puede ser una típica «generalización inductiva» de e, mientras que h_2 puede hacer aserciones, ademas de e, que no estén, en absoluto, apoyadas por e, de modo que h_2 dista de ser una generalización de e; o viceversa.

Tomemos los ejemplos siguientes.

e informa sobre un millón de cisnes observados, que son blancos, e informa de que no se han observado cisnes que no fueran blancos;

h_1 dice «todos los cisnes son blancos»;

h_2 dice «todos los cisnes en Grecia, Italia y Francia son blancos; en Inglaterra y Escandinavia, son rojos; en Asia central, son verdes; en Africa, azules; y en Australia, negros»;

b la información básica supone que no se han observado cisnes fuera de Grecia, Italia y Francia.

Ahora, haciendo que h_1 sea aplicable a todo el universo (recordemos que conjeturamos que la carga electrónica es constante en todo el universo) podemos hacerla *inicialmente* igual de probable o menos probable que h_2. (O podemos debilitar h_2 para hacerla más probable.) Podemos suponer, así, que $p\,(h_1) \leq p\,(h_2)$; y podemos suponer también que $p\,(h_1,b) \leq \leq p\,(h_2,b)$, por lo que hemos dicho sobre b. En cualquier caso, tendremos, por el teorema de multiplicación, puesto que e se sigue de h_1,b y h_2,b,

$$\frac{p\,(h_1,eb)}{p\,(h_2,eb)} = \frac{p\,(h_1,b)}{p\,(h_2,b)}$$

y así,

$$p\,(h_1,eb) < p\,(h_2,eb) \text{ si, y sólo si, } p\,(h_1,b) < p\,(h_2,b)$$

Por tanto, la evidencia e no apoya la franca generalización h_1 más de lo que apoya a la fantástica y arbitraria hipótesis h_2 (fantástica y arbitraria a la luz de la evidencia e).

Esto muestra que el argumento B) en la sección I y la relación 1) de esta sección se interpretan mal si se considera que apoyan la *inducción* o la *generalización inductiva*. Y muestra que el cálculo de probabilidades no tiene ninguna utilidad como *teoría de la inducción*.

Naturalmente, la evidencia e puede cambiar la probabilidad de una hipótesis a la que *no* sea favorable. Puede, especialmente, derribar a una hipótesis que sea incompatible con ella. Pero esto no es lo que nos interesa en este momento. (Y, aunque la evidencia reducirá a cero la probabilidad de una hipótesis a la que refute, es evidente que no necesitamos el cálculo de probabilidades para describir esta situación.)

III

Para concretar más las cosas, sean h_1 y h_2 las teorías de la gravitación de Newton y de Einstein respectivamente y sea e la evidencia disponible en 1917 (podemos incluir también la evidencia respecto al pequeñísimo movimiento del perihelio de Mercurio, que nadie consideraba entonces como un problema

serio para la teoría de Newton), de modo que tengamos que $p\,(h_1,e) = p(h_1)$. No sabemos si poner $p\,(h_1)$ mayor, menor o igual a $p\,(h_2)$; pero, mientras e pueda ser explicada por ambas teorías, la evidencia empírica e no puede, de acuerdo con 2), tener ninguna influencia sobre las probabilidades relativas.

Ahora bien, esta situación es típica: siempre existe, para toda teoría h_1 y para toda evidencia e, una teoría h_2 que está relacionada con h_1 como la teoría de Einstein lo está con la de Newton. La teoría h_2 puede no ser «concisa» o «bella» o «sencilla», pero ésa es una cuestión diferente. Siempre puede ser construida[35]. Por el mismo método podemos construir aun otra, h_3, incompatible con h_1 y también con h_2, etc. Esta consideración deja bastante claro que si hacemos que las probabilidades previas (o «absolutas») $p\,(h_1)$, $p\,(h_2)$ no sean demasiado diferentes al principio, nunca obtendremos una probabilidad tan alta como 1/2 para ninguna hipótesis, con cualquier evidencia dada. (De hecho, si empezamos, como deben hacerlo los bayesianos, como algo como $p\,(h_1) = p\,(h_2) = p\,(h_3) = \ldots$, una suposición que los bayesianos prominentes defienden por medio del principio de maximización de la «entropía probabilística», obtendremos $p\,(h_1) = 0$; pero esto no es lo que quiero señalar aquí, aunque puede utilizarse para apoyar un punto que defendí en *L. I. C.*, Apéndice *VII.) Mi argumento aquí es que nunca podemos atribuir racionalmente una probabilidad que se aproxime a 1/2 a ninguna hipótesis, dada la evidencia más excelente; la probabilidad en el sentido del cálculo de probabilidades no puede ser el instrumento que estamos buscando para explicar nuestra intuición, formulada aquí en la sección I A). Debería estar claro que nos dejamos confundir por B). Porque una probabilidad de menos de 1/2 es, naturalmente, una *improbabilidad*. Por tanto, no podemos explicar nuestra intuición inductiva por medio de un cálculo de probabilidades que nos muestra que nuestras hipótesis seguirán siendo siempre improbables y que la evidencia empírica no afecta al orden de la probabilidad de las hipótesis, excepto en aquellas hipótesis que no pueden explicar la evidencia.

Nota:

Nada de lo que se ha dicho aqui resulta afectado si introducimos en todas partes, además de e, una variable b que abarque nuestra información básica y las condiciones iniciales de modo que obtengamos

[35] Recuérdese que e es finita y, por tanto, se refiere a una gama limitada de velocidades, fuerzas, energías o lo que sea. Por tanto, sólo necesitamos elegir h_2 que, dentro de esa gama limitada, esté de acuerdo con h_1 pero se desvíe un poco de h_1 fuera de esa gama.

1') $p\ (h,be) > p\ (h,b)$

y

2') $p\ (h_1,be) < p\ (h_2,be)$ si, y sólo si, $p\ (h_1,b) < p\ (h_2,b)$

IV

Nuestro teorema 2) de la sección II muestra, pienso, que los resultados de la sección 15 son correctos. En efecto, la evidencia *e* no presta apoyo (probabilístico) *inductivo* a la hipótesis. No apoya nada dentro de la hipótesis *excepto* la parte del contenido de la hipótesis que incluye precisamente *e* y nada más. Es exactamente como lo vio Hume. *La teoría de la probabilidad no tiene en sí nada de ampliativo: no existe la inducción probabilística ampliativa.*

No obstante, la intuición formulada en la sección I, A), es válida. Si una teoría *h* ha sido bien corroborada, entonces es sumamente probable *que se parezca a la verdad. Es decir, que concuerde bien con algunos de los hechos.* Esto es sumamente probable en el sentido de que es extremadamente improbable que el éxito predictivo de una teoría poderosa y bien contrastada sea un mero accidente.

Pero no hace «probable» a *h*: porque decir que *h* es probable es decir que es *más probable que h sea verdad que lo contrario.* Esto significaría que es más probable que *h* concuerde con *todos los hechos del mundo* que lo contrario: que no existe contra-ejemplo ni hecho que la contradiga. Pero ninguna evidencia finita *e* puede decirnos nunca tal cosa[36].

[36] Véase mi *Conocimiento objetivo (Obj. Kn.)*, pp. 91-92.

CAPITULO III

OBSERVACIONES SOBRE LAS TEORIAS OBJETIVAS DE LA PROBABILIDAD

En este capítulo me propongo explicar el modo en que la teoría frecuencial de la probabilidad ha sido reemplazada por la interpretación propensivista y también por el «enfoque de la medida teórica» como se llama a menudo. Trataré de mostrar que, aunque este enfoque sustituye a la teoría frecuencial, también le da una especie de justificación *post mortem*; porque la teoría frecuencial es «*casi deducible*» del enfoque de la medida teórica. Es decir, dentro del enfoque de la medida teórica, puede mostrarse que hay una probabilidad cero de que demos accidentalmente con una sucesión aleatoria que viole las exigencias de la teoría frecuencial de von Mises. En otras palabras, es «*casi seguro*» que esas exigencias serán satisfechas[1].

La principal significación de este nuevo enfoque estriba en el hecho de que los enunciados sobre la medida teórica de la probabilidad son enunciados de probabilidad *singular:* enunciados que afirman lo que podemos llamar una «*probabilidad singular*». Sin embargo, *desde el punto de vista de la física, la mejor manera de interpretar una probabilidad singular que «casi entrañe» una frecuencia, es interpretarla como propensión física.* Así pues, la transición matemática de teoría frecuencial a teoría de la medida corresponde, sugiero, a *una transición de la interpretación estadística a la interpretación propensivista de las probabilidades físicas objetivas.*

Empiezo este capítulo con una sección que explica más plenamente la interpretación propensivista y las razones para adoptarla. Pero el capítulo consiste principalmente en observaciones históricas sobre algunos de los desarrollos, desde 1934,

---

[1] Estos resultados se deben a F. P. Cantelli (1916 y 1917) de cuya «ley fuerte de los grandes números» «casi se sigue» el axioma del límite de Venn y de von Mises (prácticamente el mismo resultado fue hallado por Harold Jeffreys y Dorothy Wrinch en 1919 y G. Pólya en 1921); y a J. L. Doob, que demostró en 1936 lo mismo para el axioma de aleatoriedad (despues de que un teorema más restringido hubiera sido demostrado por E. Hopf en 1934). Retrospectivamente, uno puede decir que ambos teoremas fueron presagiados por los teoremas de E. Borel sobre números «normales» y «enteramente normales» (1909).

de la teoría frecuencial de la probabilidad, que jugó un papel tan prominente en *L. I. C.*, sobre todo en relación con los problemas de la coherencia de la teoría de von Mises[2]. Mostraré que los supuestos de A. H. Copeland y los míos propios (que yo había tratado de reducir a un mínimo) fueron fortalecidos por A. Wald, que fue el primero que logró demostrar que era coherente una teoría casi tan fuerte como la de von Mises. La teoría de Wald fue aún más fortalecida por A. Church, que estableció en 1940, la coherencia o no-vaciedad de una clase C de colectivos que son insensibles a selecciones según *cualquier sistema de juego efectivamente calculable* (o conjunto de sistemas de juego)[3]. Este desarrollo me parece de cierta importancia, ya que establece que una teoría frecuencial muy satisfactoria puede sotenerse sobre sus propios pies y que la situación no es tal que nos fuerce a adoptar el enfoque de la medida teórica: aunque podemos adoptarlo libremente a causa de sus grandes méritos. Además es valioso saber que podemos hablar de frecuencias, incluso en sucesiones ilimitadas, sin ningún temor a la incoherencia; porque las predicciones de frecuencias —que «casi se siguen» de las hipótesis sobre la propensión— continúan siendo de una importancia decisiva para nosotros: sólo las frecuencias predichas nos permiten *contrastar* esas hipótesis.

20. EL ARGUMENTO EN FAVOR DE LAS PROPENSIONES

La interpretación subjetiva de la probabilidad puede, *quizá*, ser sostenible como interpretación de ciertas situaciones de juego —carreras de caballos, por ejemplo— en las que las condiciones objetivas del suceso están mal definidas y son irrepro-

[2] Jean Ville mostró en su *Etude Critique de la Notion de Collectif* (París, 1939), pp. 70-83, que mi teoría sobre las sucesiones libres respecto a *n* o insensibles respecto a *n* («*suites indifféréntes*») era equivalente a la teoría de los números admisibles, desarrollada por primera vez por A. H. Copeland, «Admisible Numbers in the Theory of Probability», *Am. J. of Math.* 50, 1928, y en otros muchos artículos desde entonces. Copeland opera con lo que he llamado (sección 21) «selección ordinal» en contraposición a la «selección de vecindad» que es la base de mi «insensibilidad respecto a *n*» o «libertad respecto a *n*».

[3] Simbolizando el conjunto de los «números admisibles» de Copeland por *A* (éstos son también mis propias sucesiones «absolutamente libres»); los colectivos de Wald por *W*; los colectivos de von Mises —un concepto algo vago— por *M*; y las sucesiones aleatorias de Church por *C*, encontraremos que *A*, *W*, *M* y *C* forman una sucesión decreciente de conjuntos. (No he podido encontrar ninguna referencia a Church en los últimos escritos de von Mises sobre el tema, por ejemplo, en la tercera edición alemana de su *Probability Statistics and Truth*, 1951, ni en su discusión con Doob, *Ann. Math. Statist* 12, 1941, pp. 191-217.)

ducibles. (Yo no creo realmente, sin embargo, que sea aplicable a situaciones como ésas: podría incluso establecerse un fuerte argumento —si mereciera la pena hacerlo— en favor de la concepción de que lo que un jugador o un «apostador racional» trata de hallar, para apostar, son las condiciones *objetivas*, las propensiones *objetivas*, las probabilidades *objetivas* del suceso. Así, el hombre que apuesta por caballos está ansioso por obtener más información sobre los caballos, en vez de querer informase sobre su propio estado de creencia o sobre la fuerza lógica de la información que posee.) Pero en el juego de azar típico —la ruleta o los dados o echar monedas— y en todos los experimentos físicos, la interpretación subjetiva falla por completo, como hemos visto. Porque, en todos esos casos, las probabilidades dependen de las *condiciones objetivas del experimento*.

En *L. I. C.* sólo traté de *una* interpretación objetiva de la probabilidad: la interpretación puramente estadística o frecuencial. (Considero esas dos denominaciones como sinónimas.) E intenté reconstruir esta interpretación teniendo en cuenta todas las críticas que suelen hacerse en contra suya. Aquí quiero sugerir que existe una segunda interpretación objetiva, una interpretación mejor: la interpretación propensivista. Mi sugerencia no está motivada por la creencia de que ninguna de las críticas habituales de la teoría frecuencial esté justificada. Al contrario, no dudo de la coherencia de la teoría frecuencial y creo que es posible trabajar con ella. Pero creo también que la interpretación propensivista es decididamente preferible. Algunas de mis razones para creerlos se presentarán en esta sección, mientras que otras se dejarán para más adelante.

En esta sección, la discusión se limitará únicamente al problema de interpretar la probabilidad de los «*sucesos singulares*» (u ocurrencias) y es la teoría frecuencial de la probabilidad de *sucesos singulares* lo que tengo en mente siempre que hablo aquí de la intepretación frecuencial de la probabilidad, en contraposición a la interpretación propensivista.

Se recordará que desde el punto de vista de la interpretación frecuencial de la probabilidad de un *sucueso de un cierto tipo* —tal como obtener un seis con un dado determinado— no es *nada más que* la frecuencia relativa de este tipo de suceso en una sucesión extremadamente larga (quizá infinita) de sucesos. Y si hablamos de la probabilidad de un suceso *singular* (es decir, una «ocurrencia» en el sentido de la sección 23 de *L. I. C.*; véase también la sección 71), tal como la probabilidad de obtener seis en la tercera tirada hecha después de las nueve de esta mañana con este dado, entonces, según la interpretación estadística, *sólo* queremos decir que esta tirada puede ser consi-

derada como un elemento de una sucesión de tiradas y que, en su capacidad de elemento de esta sucesión, comparte las probabilidades de esa sucesión; es decir, las probabilidades que no son *nada más que las frecuencias relativas* dentro de esa sucesión.

En esta sección me propongo argüir en contra de esta interpretación y en favor de la interpretación propensivista, haciendo uso de los resultados que se obtuvieron en el capítulo anterior respecto a *la significación de las condiciones experimentales objetivas para la interpretación de las probabilidades.* Me propongo proceder como sigue: 1) En primer lugar mostraré que, desde el punto de vista de la interpretación frecuencial hay que plantear objeciones a la interpretación propensivista que parecen hacer inaceptable a ésta última. 2) A continuación daré una respuesta preliminar a esas objeciones y después presentaré, como punto 3) una cierta dificultad a la que tiene que hacer frente la interpretación frecuencial, aunque, cuando se plantea por primera vez, no parece una dificultad seria. 4) Finalmente mostraré que, para superar esta dificultad, la interpretación frecuencial se ve forzada a adoptar una modificación que al principio parece mínima; pero la adopción de esta aparentemente mínima modificación resulta ser equivalente a la adopción de la interpretación propensivista.

1) Desde el punto de vista de una interpretación puramente estadística de la probabilidad está claro que la interpretación propensivista es inaceptable. Porque las propensiones pueden explicarse como posibilidades (o como medidas o «pesos» [ponderaciones] de posibilidades) que están dotadas de tendencias o disposiciones a realizarse y que se consideran responsables de las frecuencias estadísticas con las cuales se realizarán de hecho, en largas sucesiones de repeticiones de un experimento. Las propensiones se introducen, pues, para ayudarnos a explicar, y a predecir, las propiedades estadísticas de ciertas sucesiones; y *ésta es su única función.* Por tanto (afirmará el teórico frecuencias) no nos permiten predecir, o decir, *nada en absoluto* sobre un suceso singular, salvo que su repetición, bajo las mismas condiciones, generará una sucesión con ciertas propiedades estadísticas. Todo esto muestra que la interpretación propensivista no puede añadir nada a la interpretación frecuencial excepto una palabra nueva —«propensión»— y una nueva imagen o metáfora que se asocia con ella —la de una tendencia o desposición o impulso—. Pero estas metáforas antropomórficas o psicológicas son menos útiles incluso que las antiguas metáforas piscológicas de «fuerza» y «energía» que se convirtieron en conceptos físicos útiles sólo en la medida en que perdieron su significado original metafísico y antropomórfico.

Esta sería aproximadamente la perspectiva del teórico frecuencial. Al defender la interpretación propensivista voy a utilizar dos argumentos diferentes: una réplica preliminar 2) y un argumento que equivale a volver las tornas contra el teórico frecuencial; éste se discutirá en 3) y 4).

2) Como réplica preliminar, me inclino a aceptar la sugerencia de que hay una analogía entre la idea de propensiones y la de fuerzas, sobre todo de campos de fuerzas. Pero, aunque las etiquetas de «fuerza» o «propensión» puedan ser metáforas piscológicas o antropomórficas, la analogía importante entre ambas ideas no está ahí; está, en cambio, en el hecho de que las dos ideas llaman la atención sobre *propiedades disposicionales inobservables del mundo físico* y ayudan así a la interpretación de la teoría física. En ello está su utilidad. El concepto de fuerza —o, mejor aún, el concepto de un campo de fuerzas— introduce una entidad física disposicional, descrita por ciertas ecuaciones (más que por metáforas), para explicar aceleraciones observables. Del mismo modo, el concepto de propensión, o de campo de propensiones, introduce una propiedad disposicional de las situaciones experimentales físicas singulares —esto es, de los sucesos físicos singulares— para explicar frecuencias observables en sucesiones de repeticiones de esos sucesos. En ambos caso, la introducción de la nueva idea puede justificarse apelando a su utilidad para la teoría física. Ambos conceptos son «ocultos» en el sentido de Berkeley, o «meras palabras». Pero parte de la utilidad de esos conceptos estriba precisamente en el hecho de que sugieren que la teoría se ocupa de las probabilidades de una realidad física *inobservable* y que son sólo algunos de los efectos más superficiales de esa realidad los que podemos observar y los que nos hacen posible la contrastación de la teoría. (Véase *supra*, secciones 11 a 15.) El principal argumento en favor de la interpretación propensivista está en su poder de eliminar de la teoría cuántica ciertos elementos perturbadores de un carácter irracional y subjetivista —elementos que son más «metafísicos» que las propensiones y además, metafísicos en el mal sentido de la palabra—. Es por su éxito o su fracaso en este terreno de aplicación por lo que tendrá que ser juzgada la interpretación propensivista[4].

Una vez hecha esta réplica preliminar, procedo a mi argumento principal en favor de la interpretación propensivista. Consiste en señalar ciertas dificultades a las que debe enfrentarse la interpretación frecuencial. Llegamos así al punto 3) anunciado más arriba.

3) Se han planteado muchas objeciones contra la interpre-

[4] [Véase el tomo III del *Post Scriptum. Ed.*]

tación frecuencial de la probabilidad, especialmente en relación con la idea de sucesiones infinitas de sucesos y de límites de frecuencias relativas. No me referiré aquí a estas objeciones, porque las discutiré en la próxima sección. (En lo principal, se verá que no son válidas)[5]. Pero hay una objeción simple e importante que, que yo sepa, no se ha planteado en esta forma antes.

Supongamos que tenemos un dado cargado y que hemos quedado satisfechos, tras largas sucesiones de experimentos, respecto a que la probabilidad de sacar seis con este dado cargado se aproxima mucho a 1/4. Consideremos ahora una sucesión *b*, digamos, que consista en tiradas con este dado cargado, pero que incluya unas pocas tiradas (dos, o, quizá, tres) con un dado homogéneo y simétrico. Evidentemente, tendremos que decir, con respecto a cada una de estas pocas tiradas con el dado bueno, que la probabilidad de un seis es de 1/6 en vez de ser de 1/4, a pesar del hecho de que esas tiradas son, según nuestros supuestos, *elementos de una sucesión* de tiradas de frecuencia estadística 1/4, y a pesar del hecho de que dos o tres tiradas no pueden influir de ningún modo la frecuencia 1/4 de la sucesión larga.

Creo que esta sencilla objeción es decisiva, aunque hay varias réplicas posibles.

---

[5] Por tanto, no son las críticas habituales a la interpretación frecuencial las que me han inducido a cambiar de opinión. Pero W. C. Kneale ha sugerido lo contrario (*Observation and Interpretation*, ed. por S. Körner, 1957, p. 80). En sus comentarios sobre un extracto de esta sección (*ibid.*, pp. 66-68), hizo la siguiente suposición referente a mis razones para defender la interpretación propensivista: «Más recientemente, las dificultades de la interpretación frecuencial, es decir, los embrollos, si no las puras, contradicciones, que pueden encontrarse en von Mises, han llegado a ser bien conocidas y supongo que ésas son las consideraciones que han llevado al Profesor Popper a abandonar esa interpretación de la probabilidad». No sé de ningún «embrollo» ni «contradicción» bien conocidos en la teoría frecuencial distintos de los que discutí en *L. I. C.*, en 1934. Y todos ellos han sido aclarados hace ya mucho tiempo por Wald, Copeland, Church, yo y otros. (Véase también el próximo capítulo.) No pienso que la crítica de Kneale de la teoría frecuencial en su *Probability and Induction*, 1940, ofrezca una visión correcta de la situación lógica que ha prevalecido desde 1934. De todas formas, es cierto que hay un punto que critica Kneale (véase especialmente p. 156) que yo no discutí. Es el siguiente: en la teoría frecuencial, una probabilidad igual a 1 no significa que el suceso en cuestión vaya a ocurrir sin excepción (en las condiciones dadas). Pero este punto no constituye una debilidad de la teoría frecuencial, como afirma Kneale. Más bien, *se cumple necesariamente en toda teoría de la probabilidad adecuada que permita ser aplicada a clases infinitas.* (Esto lo mostré en *L. I. C.*, nota 14, Apéndice *VII.) Se cumple, naturalmente, en la interpretación propensivista. Todas las demás críticas mencionadas en el libro de Kneale se habían tratado en el mío, por lo que yo veo; y confío en que ninguno de sus argumentos constituya una crítica válida de la versión de Wald ni de Church ni mía, de la teoría frecuencial. Mis razones para adoptar la interpretación propensivista son, por tanto, bastante diferentes de las que supone Kneale.

Una de ellas sólo necesita ser mencionada de paso, puesto que equivale a un intento de volver a caer en la interpretación subjetivista de la probabilidad. Equivale a la aserción de que es nuestro especial *conocimiento*, la *información* especial que tenemos en relación con esas tiradas con el dado bueno, lo que cambia la probabilidad. No necesito decir que considero injustificada esta réplica, en vista de mi discusión general sobre la teoría subjetiva. Además el caso que tenemos ante nosotros sugiere otro argumento (aunque no se trata de uno muy importante) contra esta teoría subjetiva. Porque podemos no saber cuáles de las tiradas están hechas con el dado bueno, aunque podemos saber que sólo son dos o tres. En este caso, será bastante razonable apostar (siempre que estemos determinados a apostar sobre un número considerable de tiradas) sobre la base de una probabilidad muy cercana a 1/4, aunque sepamos que habrá dos o tres tiradas en las que no deberíamos aceptar apuestas en esos términos, si pudiéramos identificarlas. Sabemos que en el caso de esas tiradas, la probabilidad de un seis es menor de 1/4, que es, de hecho, 1/6; pero también sabemos que no podemos identificar esas tiradas y que su influencia debe ser muy pequeña si el número de tiradas es grande. Ahora bien, está claro que como, de todos modos, atribuimos a esas tiradas desconocidas una probabilidad de 1/6, no queremos decir, con la palabra «probabilidad», y no podemos querer decirlo, de ningún modo, «un cociente de apuesta razonable a la luz de todo nuestro conocimiento efectivo» (como quiere la teoría subjetiva). Pero dejemos a un lado la teoría subjetiva, ahora. ¿Qué puede responder el teórico frecuencial a nuestra objeción?

Como yo mismo he sido, durante muchos años, un teórico frecuencial, sé bastante bien que mi propia respuesta hubiera seguido una línea muy próxima a la que doy a continuación.

La descripción que se nos da de la sucesión *b* muestra que *b* está compuesta de tiradas con un dado cargado y, por ejemplo, de tres tiradas con un dado bueno. Estimamos, o más bien conjeturamos (sobre la base de nuestra experiencia previa o de una intuición —nunca importa cuál es la base de nuestra conjetura—), que el lado «seis» quedará hacia arriba, en una sucesión de tiradas con el dado cargado, con la frecuencia de 1/4, y en una sucesión de tiradas con el dado bueno, con la frecuencia de 1/6. Indiquemos esta última sucesión, la de las tiradas con el dado bueno, con «*c*». Entonces, nuestra información respecto a la composición de *b* nos dice i) que $p\,(a,b) = 1/4$, o muy aproximadamente, porque casi todas las tiradas son con el dado cargado y ii) que *bc* —es decir, la clase de tiradas pertenecientes tanto a *b* como a *c*— no es vacía; y

como *bc* se compone de tiradas que pertenecen a *c*, podemos afirmar que la probabilidad singular de un seis entre aquellas tiradas que pertenecen a *bc*, será de 1/6, en virtud del hecho de que esas tiradas singulares son elementos de una sucesión *c* para la cual tenemos que $p\ (a,c) = 1/6$.

Pienso que ésta hubiera sido, en conjunto, mi respuesta; y ahora me pregunto cómo hubiera podido quedarme satisfecho con una réplica de este tipo, porque ahora me parece evidente que es totalmente insatisfactoria.

Desde luego, no cabe la menor duda respecto a la compatibilidad de las dos ecuaciones (referidas a sucesiones infinitas)

i) $p\ (a,b) = 1/4$

ii) $p\ (a,bc) = 1/6$

tampoco existe duda de que esos dos casos puedan realizarse dentro de la teoría frecuencial: *podríamos* construir una sucesión *b* tal que la ecuación i) fuera satisfecha, mientras que una sucesión seleccionada *bc* —una sucesión muy larga, virtualmente infinta, cuyos elementos pertenezcan tanto a *b* como a *c*— fuera satisfecha la ecuación ii). *Pero nuestro caso no es de este tipo.* Porque *bc* no es, en nuestro caso, una sucesión virtualmente infinita. Contiene, según nuestro supuesto, exactamente tres elementos. En *bc* el seis puede no salir en absoluto, o una vez, o dos o tres veces. Pero *ciertamente* no ocurrirá con una frecuencia de 1/6 en la sucesión *bc* porque sabemos que esta sucesión contiene como mucho tres elementos.

Así pues, sólo hay dos sucesiones infinitas, o muy largas, en nuestro caso: la sucesión (efectiva) *b* y la sucesión (virtual) *c*. Las tiradas en cuestión pertenecen a ambas. Y nuestro problema es éste: aunque pertenezcan a ambas sucesiones y aunque sólo sepamos que esas tiradas concretas *bc* ocurren en algún punto de *b* (no nos dicen dónde y, por tanto, no somos capaces de identificarlas), no tenemos ninguna duda de que, en su caso, la probabilidad singular apropiada, la verdadera, es 1/6 en vez de 1/4. O, dicho de otro modo, aunque pertenecen a ambas sucesiones, no tenemos duda de que su probabilidad singular ha de estimarse como igual a la frecuencia de la sucesión *c* y no *b*, simplemente porque son tiradas con un dado diferente (bueno) y porque estimamos o conjeturamos que, en una sucesión de tiradas con un dado bueno, el seis saldrá en 1/6 de los casos.

4) Todo esto significa que el teórico frecuencial se ve forzado a introducir una modificación en su teoría, aparentemente una modificación muy ligera. Dirá ahora que una sucesión admisible de sucesos (una sucesión de referencia, un «colec-

tivo») debe ser siempre una sucesión de condiciones repetidas. O, más generalmente, dirá que las sucesiones admisibles deben ser sucesiones efectivas o sucesiones virtuales que se *caractericen por un conjunto de condiciones generadoras*, por un conjunto de condiciones cuya repetida realización produce los elementos de una sucesión independiente.

Si se introduce esta modificación, entonces nuestro problema queda resuelto al instante. Porque la sucesión *b* ya no será una sucesión de referencia admisible. La parte de ella (es decir, $b\bar{c}$) que consiste en tiradas con el dado cargado, constituirá una sucesión admisible y no se plantean cuestiones sobre ella. La otra parte, *bc*, se compone de tiradas con un dado normal y pertenece a una sucesión virtual *c* —también admisible— de tales tiradas. Tampoco hay problema aquí. Está claro que una vez que se ha adoptado la modificación, la interpretación frecuencial no tiene ya ninguna dificultad.

Además parece que lo que he descrito como una «modificación» sólo enuncia explícitamente un supuesto que la mayoría de los teóricos frecuenciales (incluyéndome a mí mismo) han dado siempre por sentado.

Pero, si miramos más atentamente esta aparentemente mínima modificación, encontramos que equivale a una transición de la interpretación frecuencial a la interpretación propensivista.

La interpretación frecuencial siempre considera la probabilidad como relativa respecto a una sucesión que se supone «dada»: y opera con el supuesto de que una probabilidad es *una propiedad de una sucesión dada*. Pero con nuestra modificación, la sucesión es definida a su vez por su conjunto de *condiciones generadoras;* y de tal forma que puede decirse ahora que la probabilidad es *una propiedad de las condiciones generadoras*.

Pero esto supone una diferencia muy grande, especialmente respecto a la probabilidad de un suceso singular (o una «ocurrencia»). Porque podemos decir ahora que el suceso singular *a* tiene una probabilidad $p\,(a,b)$ debido al hecho de que es un suceso producido, o seleccionado, de acuerdo con las condiciones generadoras $b\bar{c}$, más que debido al hecho de que es un elemento de una sucesión *b*. De este modo, un suceso singular puede tener una probabilidad incluso aunque sólo ocurra una vez; porque su probabilidad es una propiedad de sus condiciones generadoras. Es generada por ellas.

Hay que admitir que el teórico frecuencial todavía puede decir que la probabilidad, aunque sea una propiedad de las condiciones generadoras, es igual a la frecuencia relativa dentro de una sucesión virtual o efectiva generada por esas condiciones. Pero si pensamos en esto más atentamente, resulta bas-

tante claro que nuestro teórico frecuencial se ha convertido, sin advertirlo, en un teórico propensivista. Porque si la probabilidad es una propiedad de las condiciones generadoras (digamos que la disposición experimental) y se considera, por tanto, dependiente de esas condiciones, entonces la respuesta del teórico frecuencial supone que la frecuencia virtual debe depender también de esas condiciones. Pero esto significa que tenemos que imaginarnos las condiciones como dotadas de una tendencia o disposición o propensión a producir sucesiones con frecuencias iguales a las probabilidades; que es precisamente lo que afirma la interpretación propensivista.

Podría pensarse que es posible evitar el último paso —la atribución de propensiones a las condiciones generadoras— hablando de meras posibilidades en vez de hablar de propensiones. De este modo, uno puede tener la esperanza de evitar lo que parece ser el aspecto más cuestionable de la interpretación propensivista: su similaridad intuitiva con «fuerzas vitales» y otros antropomorfismos a los que tan a menudo se ha calificado de pseudo-explicaciones estériles.

La interpretación de las probabilidades en términos de posibilidades es, desde luego, muy antigua. Podemos, en pro del argumento, suprimir las conocidas objeciones (ejemplificadas por el caso del dado cargado) contra la definición clásica de la probabilidad en términos de posibilidades *iguales*, como el número de posibilidades favorables dividido entre el número de todas las posibilidades; y podemos limitarnos a casos como los de dados o peniques simétricos, para ver cómo se compara esta definición con la intepretación propensivista.

Las dos interpretaciones tienen mucho en común. Ambas se refieren primordialmente a sucesos singulares y a las posibilidades inherentes en las condiciones bajo las cuales tiene lugar cada suceso. Y ambas consideran esas condiciones como reproducibles en principio, de modo que pueden dar lugar a una sucesión de sucesos. La diferencia, al parecer, estriba no solamente en esto: una interpretación introduce esas cuestionables propensiones metafísicas, mientras que la otra se refiere simplemente a las simetrías de las condiciones, a las posibilidades iguales que las condiciones dejan abiertas.

Con todo, este acuerdo es sólo aparente. No es difícil ver que las simples posibilidades son inadecuadas a nuestro propósito —o al del físico o al del jugador— y que incluso la definición clásica supone, implícitamente, que a las posibilidades iguales hay que atribuirles iguales *disposiciones o tendencias o propensiones a realizar las posibilidades en cuestión.*

Esto puede mostrarse muy fácilmente si consideramos primero unas equi-posibilidades muy próximas a cero. Un ejemplo

de una equi-posibilidad muy próxima a cero sería la probabilidad de cualquier sucesión definida de ceros y de unos de la longitud n: Existen 2^n de tales sucesiones, de tal modo que en el caso de equi-posibilidad, cada posibilidad tiene el valor de $1/2^n$ que, para un valor muy grande de n es muy próximo a cero. La posibilidad complementaria es, naturalmente, igual de próxima a uno. Ahora bien, esas posibilidades cercanas a cero se interpretan generalmente como «casi imposibles» o como que «casi nunca se realizan», mientras que, naturalmente, las posibilidades complementarias, que son próximas a uno, se interpretan como «casi necesarias» o como que «casi siempre se realizan».

Pero si se admite que las posibilidades cercanas a cero y cercanas a uno deben interpretarse como predicciones —«no ocurren casi nunca» y «ocurren casi siempre»— entonces puede mostrarse fácilmente que las dos posibilidades de obtener cara o cruz, que se suponen exhaustivas, exclusivas e iguales, deben interpretarse también como predicciones. Corresponden a la predicción «es casi seguro que se realizarán, a la larga, aproximadamente en la mitad de los casos». Porque podemos mostrar, con ayuda del teorema de Bernoulli (el ejemplo de las sucesiones de longitud n) que esta interpretación de posibilidades de 1/2 es *lógicamente equivalente* a la interpretación que acabamos de dar, de posibilidades cercanas a cero o a uno.

Para expresar lo mismo en forma algo diferente, las meras posibilidades no pueden dar lugar a ninguna predicción. Es posible, por ejemplo, que un terremoto destruya mañana *todas* las casas entre los paralelos 13 norte y 13 sur (y *ninguna* otra casa). Nadie puede calcular esta posibilidad, pero la mayoría la consideraría muy pequeña; y mientras la pura posibilidad como tal no da lugar a ninguna predicción, la estimación de que es extremadamente pequeña puede convertirse en base de la predicción de que el suceso descrito no tendrá lugar («con toda probabilidad»).

Así, la estimación de la *medida* de una posibilidad —esto es, la estimación de la probabilidad que se le atribuye— siempre tiene un aspecto predictivo, mientras que no predeciríamos un suceso si nos dijeran solamente que ese suceso es posible. En otras palabras, no suponemos que la posibilidad como tal tenga una tendencia a realizarse; pero interpretamos las medidas de la probabilidad o las «ponderaciones» que se atribuyen a la posibilidad como medidas de su disposición o tendencia o propensión a realizarse; y en física (o en apuestas) estamos interesados en medidas o «ponderaciones» de posibilidades como licencia para hacer predicciones. Por tanto, no podemos escapar al hecho de que tratamos las medidas de las posibilidades como disposiciones o tendencias o propensiones. Mi razón para esco-

ger la denominación «*interpretación propensivista*» es que quiero destacar este punto que, como muestra la historia de la teoría de la probabilidad, pasa fácilmente desapercibido.

Por ello no me intimida la alegación de que la propensión es un concepción antropomórfica o de que es similar a la concepción de una fuerza vital. (Esta concepción ha sido efectivamente estéril hasta el momento y parece cuestionable. Pero la disposición o tendencia o propensión de la mayoría de los organismos a luchar por la supervivencia no es una concepción estéril, sino muy útil; y la esterilidad de la idea de fuerza vital parece debida al hecho de que promete añadir, y no lo hace, algo importante a la aserción de que la mayor parte de los organismos muestran una propensión a luchar por la existencia y, al hacerlo, desarrollan otras propensiones, como la de investigar su entorno y la de ocupar nuevos nichos ecológicos.)

En resumen, la interpretación propensivista puede presentarse como una interpretación que conserva la concepción de que las probabilidades son frecuencias estadísticas conjeturadas o estimadas, en sucesiones largas (efectivas o virtuales). Pero al llamar la atención sobre el hecho de que esas sucesiones están definidas por la manera en que se generan sus elementos —es decir, por las condiciones generadoras— podemos mostrar que tenemos que atribuir inevitablemente nuestras probabilidades conjeturadas a esas condiciones generadoras: tenemos que admitir que dependen de esas condiciones y que pueden cambiar con ellas. Esta modificación de la interpretación frecuencial lleva casi inevitablemente a la conjetura de que las probabilidades son propiedades disposicionales de esas condiciones: es decir, propensiones. Esto nos permite interpretar la probabilidad de un suceso *singular* como una propiedad del suceso singular en sí, a medir por una frecuencia estadística *potencial o virtual* conjeturada y no por una frecuencia *efectiva* u observada.

Como todas las propiedades disposicionales, las propensiones exhiben una cierta similaridad con las potencialidades aristotélicas. Pero hay una diferencia importante: no pueden estar inherentes en las *cosas* individuales, como se inclinarían a pensar los aristotélicos. No son propiedades inherentes al dado o al penique, sino a algo un poco más abstracto, aunque físicamente real: son propiedades relacionales de la situación objetiva total; propiedades ocultas de una situación cuya dependencia precisa de la situación sólo podemos conjeturar. Y si queremos contrastar nuestra conjetura, tenemos que tratar de mantener constante la *situación pertinente*, manteniendo constantes algunas de las condiciones en cada repetición del suceso. En este aspecto, las propensiones se asemejan a las fuerzas o a los campos de fuerzas: una fuerza newtoniana no es una propiedad de una

cosa, sino una propiedad relacional de al menos dos cosas; y las fuerzas efectivas resultantes en un sistema físico son siempre una propiedad de todo el sistema físico. Fuerza, como propensión, es un concepto relacional.

Este aspecto relacional de una propensión (digamos que de una situación experimental) pasa fácilmente desapercibido: podemos pensar que la propensión a que salga cara o cruz en la mitad de las tiradas es una propiedad inherente del penique. Pero totalmente aparte del hecho de que no es una propiedad del penique, sino del *tirar* un penique, descubriremos que obtenemos probabilidades más bajas si dejamos caer el penique sobre una superficie de arena blanda o de barro (donde puede quedar en reposo de canto) en vez de una pista de tenis; lo que muestra que puede haber diversas condiciones experimentales a tener en cuenta incluso en los casos más sencillos.

Estos resultados apoyan los resultados de nuestro análisis del papel de *b* —el segundo argumento— en «*p (a,b)*», y son apoyados por ellos; y muestran que, aunque podemos interpretar «*b*» como el nombre de una sucesión (potencial o virtual) de sucesos, no debemos admitir toda sucesión posible: sólo deben admitirse aquellas sucesiones que puedan describirse como repeticiones de una situación que engendra ciertos resultados posibles y que puede caracterizarse por el método de su generación, es decir, por un conjunto generador de condiciones experimentales.

Hay una posibilidad de malinterpretar mis argumentos y, sobre todo, los de esta sección. Porque quizá puedan tomarse como ilustraciones del método de *análisis de significados:* lo que he hecho o intentado hacer, podría argüirse, es mostrar que la palabra «probabilidad» se usa, en ciertos contextos, para denotar propensiones. Quizá incluso yo he estimulado esta mala interpretación, especialmente en esta sección, al sugerir que la teoría frecuencial es, en parte, el resultado de un análisis de significados equivocados o incompleto. Sin embargo, no sugiero que se ponga en su lugar otro análisis de significados. Esto se verá claramente en cuanto se entienda que lo que propongo es *una nueva hipótesis física* (o quizá una hipótesis metafísica), análoga a la hipótesis de las fuerzas newtonianas. Es la hipótesis según la cual toda disposición experimental (y, por tanto, *todo estado de un sistema*) genera propensiones que a veces pueden ser contrastadas por medio de frecuencias. Esta hipótesis es contrastable y está corroborada por ciertos experimentos cuánticos. El experimento de las dos ranuras, por ejemplo (*cf.* tomo III del *Post Scriptum,* sección 18), puede decirse que es algo parecido a un experimento crucial entre la interpretación puramente estadística de la probabilidad y la interpreta-

ción propensivista y puede decirse que decide la cuestión contra la interpretación puramente estadística.

La interpretación propensivista se discutirá más plenamente en los restantes tomos del *Post Scriptum*. Habrá que juzgarla a la luz de esas discusiones.

En este capítulo mostraré que la transición de la interpretación frecuencial a la propensivista corresponde a la transición de la *teoría frecuencial matemática*, desarrollada por von Mises, Copeland, Wald, Church (y yo), al *tratamiento neoclásico o de la medida teórica de la probabilidad*, que estoy convencido de que es mejor que la teoría frecuencial, no sólo desde un punto de vista filosófico, sino también desde uno puramente matemático. Al final del tomo II del *Post Scriptum*, que está dedicado al problema del determinismo, me propongo mostrar que lo que ha impedido durante tanto tiempo la aceptación consciente de la interpretación propensivista ha sido la creencia en el determinismo metafísico. En el tomo III del *Post Scriptum*, sobre la teoría cuántica, se pondrá a prueba la utilidad de la interpretación propensivista. En mi epílogo (también en el tomo III) me propongo mostrar que, con ayuda de la interpretación propensivista, puede construirse una nueva metafísica de la física, un nuevo programa de investigación para la física, que unifique la mayoría de sus programas más antiguos y que, además, parece ofrecer posibilidades para una unificación de las ciencias física y biológica.

21. DONDE TIENE EXITO LA TEORIA FRECUENCIAL

Cualquiera que sea la interpretación que adoptemos de los enunciados científicos de la probabilidad, no hay duda de que la interpretación frecuencial sigue siendo de fundamental importancia, puesto que son siempre enunciados de frecuencias lo que sometemos a contrastaciones empíricas. Por esta razón, empezaré aquí con los problemas que traté en el capítulo más largo (capítulo 7) de *L. I. C.*

Cuando escribí aquel capítulo, la discusión del llamado axioma o postulado de von Mises de *aleatoriedad* (o de la futilidad de los sistemas de juego), estaba aproximadamente en su punto culminante.

Este punto culminante lo describe Karl Menger[6] como sigue: «En aquella época ocurrió un segundo acontecimiento

[6] Véase K. Menger, «The Formative Years of Abraham Wald...» una contribución al 23.° volumen de los *Annals of Mathem. Statistics*, 1952, que estaba dedicado a la memoria de Abraham Wald.

que resultó de crucial importancia para la vida y la obra de Wald. El filósofo vienés Karl Popper... trató de precisar la idea de una sucesión aleatoria y remediar los defectos obvios de la definición de colectivos de von Mises. Despues de haber oído (en el Círculo Filosófico de Schlick)[7] una exposición semitécnica de las ideas de Popper, pedí a éste que presentara el importante tema con todo detalle en el Colloquio Matemático. Wald se interesó enormemente[8] y el resultado fue su magistral artículo sobre la auto-coherencia de la noción de colectivos... en *Ergebnisse...* Fue a través de este artículo sobre colectivos y un estudio de series temporales... emprendido por sugerencia de Morgenstern[9] como Wald se interesó por los fundamentos de la estadística»[10].

Espero que se me perdone el haber referido este incidente. No sólo fue importante para Wald, sino también para mí. Porque el trabajo de Wald era una generalización del mío, de una amplitud y una profundidad que iban mucho más allá de lo que yo podía aspirar en este terreno. (En consecuencia, no llegué a publicar ciertos artículos menores sobre el tema, que había preparado y a los que había aludido en *L. I. C.*) El método de Wald puede describirse, aproximadamente, como el empleo de una especie de «argumento diagonal» que nos permite construir sucesiones insensibles a cualquier conjunto numerable de sistemas de juego.

Los resultados de Wald eran, brevemente, éstos: Wald mostró que, dado cualquier conjunto numerable de sistemas de juego, existen colectivos —de hecho, todo un continuo de colectivos— que son insensibles a todos esos sistemas de juego (en otras palabras, existen sucesiones con frecuencias convergentes que los sistemas de juego no logran afectar). Mostró, además, que si el conjunto de sistemas de juego «es definido de una manera constructiva», entonces uno puede *construir efecti-*

---

[7] Yo no estaba presente cuando tuvo lugar esta exposición de mis ideas, ya que no era miembro del «Círculo de Viena» de Schlick. (Todas las notas a pie de página en el pasaje citado de Menger han sido añadidas por mí. K. R. P.)

[8] En realidad, Wald leyó mi libro, a instancias de Menger, unos dos meses antes del 84.° Coloquio en el que tanto él como yo leímos nuestros artículos; y él tenía ya preparados todos sus principales resultados para entonces. *Cf.* la nota 10, más adelante.

[9] O. Morgenstern era entonces director del Instituto para la Investigación del Ciclo Económico en Viena.

[10] Un largo artículo de Abraham Wald en *Ergebnisse eines Mathematischen Kolloquiums*, editado por Karl Menger en el núm. 8 (1937); pero él había presentado sus resultados en febrero de 1935, en los Kolloquiums 84 y 85; véase *Ergebnisse*, núm. 7, 1936, p. 12. Un resultado final (correspondiente a la Parte 3, pp. 70-73 del largo artículo y descrito aquí en la próxima nota) fue añadido en el Kolloquium 86, el 1 de marzo de 1935.

vamente ejemplos de colectivos que sean insensibles a todos esos sistemas de juego[11].

El resultado de Wald[12] recibió los últimos toques, por así decirlo, en un artículo de Alonzo Church. Wald había utilizado las palabras «definido de una manera constructiva» y «puede construirse efectivamente» en un sentido no técnico: había afirmado, sencillamente, que siempre que se nos presentase un conjunto (numerable) de sistemas de juego con ayuda de alguna indicación o método para construir todos los sistemas pertenecientes al conjunto, entonces esta indicación o método podría utilizarse para construir colectivos que fueran insensibles a una selección según todos los sistemas de juego del conjunto.

Ahora Church señalaba[13] que éste era un caso al que podía aplicarse el concepto de *calculabilidad efectiva*, para el cual, él había propuesto una definición formal en 1936. Recordó que von Mises había criticado severamente[14] a todos aquellos que,

---

[11] Un resultado secundario, aunque muy sorprendente, de Wald (véase la nota anterior) fue el siguiente. Si damos a la «selección por vecindad» un sentido algo diferente del que se le dio en *L. I. C.* (en donde depende solamente del elemento cuya selección hay que decidir), cortando la sucesión en segmentos que no se superpongan, de una longitud definida y considerando que cada uno de esos segmentos constituye un «vecindario» de sus elementos, entonces los colectivos o las sucesiones bernoullianas (incluidas las mías) *no* son insensibles a las «selecciones por vecindad» en este sentido especial. Véanse las notas *4 y *5 a la sección 58 y *3 a la sección 60 de *L. I. C.*

[12] El resultado de Wald era, esencialmente, una generalización de resultados alcanzados previamente por A. H. Copeland. Una generalización afín fue lograda casi al mismo tiempo por J. L. Doob, «Note on Probability», *Annals of Mathematics* (Segunda Serie) 37, pp. 363-367 (publicado en abril de 1936; recibido el 16 de septiembre de 1935). La primera publicación breve por Wald de sus resultados había sido un poco antes (*Comptes rendus de l'Académie des Sciences*, París, tomo 202, pp. 180-183, 20 de enero de 1936); y Wald leyó su artículo por primera vez en febrero de 1935, como he mencionado en la nota 10. Véase también A. H. Copeland, «Consistency of the Conditions Determining Collectives», *Trans. Am. Math. Soc.* 42, 1937, pp. 333 ss., y W. Feller, «*Ueber die Existenz vo sogenannten Kollektiven*», en *Fundam. Mathem.* 32, 1939, pp. 87 ss.

[13] Alonzo Church, «On the Concept of a Random Sequence», *Bull. Am. Math. Soc.* 46, 1940, pp. 130-135.

[14] Véase especialmente *Probability Statistics and Truth* (1939), p. 136 (2.ª edición alemana, pp. 105 ss.) A. H. Copeland había tratado ya en 1928 (*Am. J. of Math.* 50, 1928; 53, 1931; etc.) los «números admisibles», que eran idénticos a lo que Reichenbach llamó más tarde «sucesiones normales». Son sucesiones insensibles a la «selección ordinal normal»; y Copeland había demostrado la existencia de tales números, pero sin construir un ejemplo. Yo, por otro lado, había partido de un requisito que me parecía intuitivamente más importante: la insensibilidad a *las selecciones según los predecesores.* Demostré que este requisito era suficiente para establecer la insensibilidad tanto a las selecciones «ordinales normales» como a las de «pura vecindad»; y poco después, Ville demostró la equivalencia de la «selección ordinal normal» con mi requisito. También di un método de construcción de sucesiones insensibles a las selecciones según *n* predecesores.

como Copeland, habían operado con colectivos[15] que *pudieran* construirse con ayuda de una regla; e incluso más severamente a aquellos que, como yo, habían propuesto métodos para tales construcciones[16]. Para tales sucesiones tienen que existir siempre sistemas de juego que tengan éxito, puesto que las sucesiones están sujetas a construcción matemática; y von Mises señaló que siempre puede haber muchos más sistemas de juego que tengan éxito, para esas sucesiones.

Por mi parte, yo nunca consideraré que esto fuera una objeción seria. Lo único que yo quería era derivar el formalismo de la teoría de la probabilidad del *supuesto de que la probabilidad del enésimo elemento de la sucesión era independiente de las propiedades de todos sus predecesores.* Church, sin embargo, encontraba importantes las objeciones de von Mises. Replicó a ellas señalando que *cualquier* sistema practicable de selección (sistema de juego) debe ser tal que nos permita *calcular efectivamente* los elementos que deben ser seleccionados (para apostar por ellos). Propuso, por tanto, definir una *sucesión aleatoria* dejando inalterada la primera condición de von Mises («axioma de convergencia») y cambiando la segunda condición de von Mises («axioma de aleatoriedad» o «axioma de exclusión de los sitemas de juego») de tal modo que se exigiese de todas las sucesiones aleatorias *insensibilidad a todas las funciones de selección efectivamente calculables.*

Al excluir así todos los sistemas de juego efectivamente calculables, es decir, practicables (todos los que son susceptibles de recibir una formulación matemática precisa) Church, en mi opinión, ha logrado caracterizar precisamente el tipo de colectivo que tenía en mente von Mises. Church mostró que la prueba de Wald era aplicable a este caso[17]. Así pues, la existencia de colectivos —o «sucesiones aleatorias», como las llamó Church— quedó demostrada.

Estos resultados me parecen establecer la justificación más completa de la teoría frecuencial de von Mises que es razonable

---

[15] El término de Copeland es «números admisibles».

[16] Wald creía (*cf. Ergebnisse, op. cit.*, p. 44) que el mío era el primer método para construir sucesiones que fueran realmente insensibles a las selecciones según los predecesores. Yo también creía que era el primero (véase el Apéndice IV); pero más tarde supe por von Mises que él mismo había dado un método para construir sucesiones bernoullianas en 1933. (*Math. Annalen* 108, p. 769.) Ni Wald ni yo sabíamos esto. No obstante, mis sucesiones (*cf.* la nueva nota 2 a la sección 55 y Apéndice IV) resultan ser diferentes en varios aspectos a las de von Mises; en particular, se hacen más pronto insensibles respecto a *n* para un *n* dado, lo que parece hacer de ellas, desde el principio, mejores copias de sucesiones aleatorias empíricas. (*Cf.* Apéndice *VI.)

[17] Esto se sigue del resultado de Wald en vista del hecho de que el conjunto de funciones efectivamente calculables es numerable (no efectivamente).

exigir. Deberían silenciar a todos los críticos. Incluso los que (como yo mismo) ponían objeciones al «axioma de convergencia» (o «axioma del límite»), recibieron una respuesta efectiva.

Porque Church mostró que un importantísimo resultado de Borel[18] podía extenderse a las «sucesiones aleatorias» de Church, o, al menos, a las sucesiones con dos propiedades, «0» y «1», siempre que tuvieran igual distribución, es decir, siempre que $p\ (0) = p\ (1) = 1/2$[19].

El resultado de Church que tengo en mente es éste. Si consideramos todas las infinitas «alternativas» posibles —es decir, todas las infinitas sucesiones posibles de ceros y unos— entonces, *casi todas* ellas son sucesiones aleatorias en el sentido de Church.

Esto supone, evidentemente, primero, que todas las sucesiones alternativas, si se continúan para siempre, son convergentes; y en segundo lugar, que casi todas ellas tienen un carácter «fortuito» o «aleatorio» y, aún más, supone que tales sucesiones aleatorias *existen.*

Con «casi todas» quiero decir aquí los siguiente: interpretemos las sucesiones de ceros y unos como desarrollos fraccionarios binarios de los números reales entre cero y uno. Se ve entonces que las que no son aleatorias, es decir, las que no cumplen las dos condiciones de aleatoriedad de Church, forman un conjunto de medida cero.

Para entender intuitivamente este teorema, para alternativas con probabilidad 1/2, podemos considerar todas las sucesiones posibles de longitud 2, 4, 6, ... ordenadas según su magnitud. (Por razones de espacio, sólo doy los dos primeros conjuntos.)

Longitud 2: Número de sucesiones: 4

00	10
01	11

[18] E. Borel, *Leçons s. l. Théorie des Fonctions* (ed. 1914, 1928), nota v. (Copeland había usado anteriormente el resultado de Borel en relación con sus propios «números admisibles».) Borel publicó por primera vez sus resultados en su artículo *Les probabilités dénombrables et leurs applications arithmétiques. Rend Palermo* 27, 1909.

[19] Con $p\ (0)$ y $p\ (1)$ simboliza la probabilidad (absoluta) de la ocurrencia de 0 ó 1; del mismo modo, con $p\ (0,a)$ y $p\ (1,a)$, simbolizo las correspondientes probabilidades relativas dado a. Church asevera el teorema sólo para alternativas con probabilidad 1/2. Sin embargo, los resultados de Borel indican un teorema más general, como ha señalado Copeland (véase, por ejemplo, su artículo en *Erkenntnis* 6, 1936, pp. 189-203. Véase también el último párrafo de J. L. Doob, «Note on Probability», *Ann. of Maths.*, Segunda Serie, 37, 1936, pp. 363-367.)

Longitud 4: Número de sucesiones: 16

0000	1000
0001	1001
0010	1010
0011	1011
0100	1100
0101	1101
0110	1110
0111	1111

En general, el número de sucesiones diferentes de longitud n es 2^n.

Como advirtió Bernoulli hace mucho tiempo, se cumple lo siguiente:

1) La frecuencia relativa de las sucesiones que tienen *exactamente* igual distribución decrece con su longitud. (Es de 1/2 en las sucesiones de longitud 2, de 3/8 en las de longitud 4, de 5/16 en las de longitud 6.)

2) No obstante, la frecuencia relativa de las que tienen una distribución casi igual, aumenta. (Esto no puede ilustrarse muy bien si no procedemos, al menos, hasta sucesiones de longitud 6.)

3) La frecuencia relativa de las sucesiones que tienen un carácter intuitivamente aleatoriforme y que son, con buena aproximación, insensibles a la «selección ordinal normal», aumenta con la longitud de las sucesiones. (Esto sólo puede ilustrarse adecuadamente si consideramos sucesiones de mayor longitud, al menos hasta 6 y 8.)

Ahora bien, si las sucesiones se hacen cada vez más largas, una proporción mayor se desviará muy poco de $p\,(0) = p\,(1) = 1/2$; y una proporción mayor se hará casi insensible cada vez a más métodos de selección. De este modo, obtenemos nuestro teorema.

Consideremos ahora la crítica que suele hacerse al axioma de convergencia, a la luz de este resultado. Normalmente se ha dicho:

a) que es inútil postular de una sucesión que no puede calcularse que converge, de acuerdo con una regla matemática;

b) que según sea la teoría de la probabilidad, todas las sucesiones son posibles y compatibles con cualquier suposición probabilística y, por tanto, también una sucesión como la siguiente[20]:

01 0011 00000011111 0000000000000000001111111111111111

[20] En esta sucesión, $a_1 = 0$, $a_2 = 1$, y cada bloque de ceros va seguido por un bloque de unos de la misma longitud, seguido a su vez por un bloque de

que oscila entre las frecuencias 1/2 y 1/3 y, por tanto, no es convergente. Pero es inadmisible excluir esta sucesión divergente, puesto que, evidentemente, es una sucesión posible.

Ahora puede darse respuesta a esas objeciones:

a) Casi todas las sucesiones que *no* son efectivamente calculables de acuerdo con una regla matemática, tienen frecuencias convergentes.

b) Aunque existen sucesiones divergentes, pueden despreciarse, por tanto. Esto nos da una teoría que idealiza muy ligeramente la situación y la simplifica.

Visto a esta luz, el «axioma de convergencia» (o «axioma del límite», como lo llamó von Mises) pierde totalmente su carácter aparentemente cuestionable. Porque en lugar de establecerlo como un postulado arbitrario, ahora adopta el carácter de una versión idealizada o simplificada de un teorema. Y en lugar de parecer aplicable a sucesiones infinitas no existentes empíricamente, puede ahora considerarse que idealiza simplemente una propiedad compartida por *casi todas* las sucesiones *finitas* de gran longitud y cada vez por más sucesiones finitas al aumentar su longitud.

Todo esto es muy claro y satisfactorio. Sólo puedo repetir que, en mi opinión, justifica completamente el enfoque frecuencial.

Sin embargo, esta misma justificación del enfoque frecuencial lo sustituye: la teoría frecuencial se hace obsoleta en el mismo momento en el que puede justificarse por completo matemáticamente. Porque la teoría que la justifica no es, a su vez, una teoría frecuencial en el sentido de von Mises: es, esencialmente, una teoría que mide posibilidades o conjuntos de posibilidades, como la teoría clásica (originalmente, creación de Bernoulli). Quizá pueda llamarse teoría «neoclásica». Por su capacidad de justificar la teoría frecuencial demuestra ser la más fuerte de las dos; en realidad hace superflua la teoría frecuencial. En otras palabras, una vez que se ha realizado el objetivo de Bernoulli, y se ha tenido un puente sobre la brecha que separaba la teoría clásica o, más bien, la neoclásica, de la teoría frecuencial, ésta última pierde, por medio de su justificación, su existencia independiente y pasa a formar parte de la primera.

---

ceros cuya longitud es igual al doble del número de ceros en todas las ocurrencias anteriores.

22. DONDE FRACASA LA TEORIA FRECUENCIAL

La teoría frecuencial no *necesita* de la teoría neoclásica: es bastante autónoma. Y, sin embargo, falla porque no es suficientemente general. Hay problemas, y soluciones, del mayor interés que no pueden plantearse dentro del ámbito de la teoría frecuencial.

Para mostrar esto con ayuda de un ejemplo, expondré con más detalle una forma simple del teorema mencionado en la sección anterior, según el cual, en *casi todas* las alternativas, la frecuencia relativa de unos tiene un límite.

Sea a una alternativa (es decir, una sucesión de ceros y unos). Sea n' el número de unos que aparecen hasta el enésimo lugar de a, de forma que n'/n sea la frecuencia relativa de los unos hasta en enésimo lugar de a.

Un matemático dirá que n'/n tiene un límite —podemos llamarlo $p\,(l,a)$— si, y sólo si, para cada pequeña fracción ε, que elegimos tan pequeña como deseemos, pero con $\varepsilon > 0$, existe un número m tal que, del emésimo lugar en adelante, n'/n se desvíe de $p\,(l,a)$ en menos de ε; o, en símbolos:

(*) Existe un m tal que, para todo n elegido de tal forma que $n > m$,

$$|n'/n - p\,(l,a)| < \varepsilon$$

Ahora, para una alternativa empírica *aleatoria* a (para un colectivo), nunca podemos calcular efectivamente el número m correspondiente a cualquier ε elegida (simplemente porque el colectivo no está determinado por una regla matemática). Lo que podemos hacer, sin embargo, es esto: podemos calcular efectivamente un m para toda ε elegida, por pequeña que sea, tal que la *probabilidad de que (*) falle* sea menor que ε; o, lo mismo, que *la probabilidad de que (*) se cumpla* sea mayor que $1 - \varepsilon$. De hecho, es más que suficiente hacer a m igual a $1/\varepsilon^3$, que será un número muy grande si ε es una fracción pequeña. [Si $\varepsilon = 1/1.000$, entonces $m = 1.000$ millones; lo que significa que hay una probabilidad menor que 1/1.000 de que después del millonésimo lugar de una alternativa vuelva a ocurrir una desviación de la frecuencia relativa de $p\,(l,a)$ que exceda 1/1.000][21].

[21] El corolario simplificado del teorema de Cantelli usado aquí en el texto

1) $$m > 1/\varepsilon^3$$

es válido bajo la condición de que $\varepsilon \leq 0{,}037$ [si tomamos 3), *infra*, como nuestra mejor estimación de m]. Esto entraña que $m = 19.742$. La fórmula 1)

De esta forma, el teorema puede interpretarse dentro de la teoría frecuencial, como lo ha mostrado von Mises[22], como sigue.

Tomamos la alternativa *a* y formamos a partir de ella un nuevo colectivo *b* del modo siguiente: cortamos *a* en segmentos muy largos (que no se superpongan) de una longitud elegida *n*, donde $n > m$. Los elementos del nuevo colectivo *b* son esos largos segmentos de *a*. Lo que el teorema afirma es que dentro de los segmentos que son elementos de *b*, una desviación de *p (l,a)* que sea mayor de 1/1.000, no se encontrará, por término medio, más que una vez en cada milésimo segmento; de modo que la frecuencia relativa de esas desviaciones tiende, dentro del colectivo *b* a un límite que es menor que 1/1.000 —sin importar lo grande que hayamos escogido a *n*—.

No puede oponerse ninguna objeción a esta interpretación del teorema.

Sin embargo, el teorema expuesto puede convertirse en la base de un corolario que afirma que *casi todas las alternativas tienen frecuencias convergentes.* Y el razonamiento que lleva desde el teorema a este corolario no puede reproducirse dentro de la teoría frecuencial.

Dentro del marco de la teoría clásica puede darse al teo-

---

puede obtenerse a partir de la 2) y también de la versión del teorema de Cantelli que discute von Mises (véase la próxima nota) y que puede escribirse

$$m > 1/\varepsilon^2\eta$$

Podemos demostrar 1) poniendo $\eta = \varepsilon$ o bien en esta versión o bien en el siguiente obvio, y sin embargo, interesante corolario del teorema de Cantelli (cuyas condiciones entrañan $m \geq 16.166$):

2) $$m = (\varepsilon + \eta)/2\varepsilon^3\eta$$

(siempre que $\varepsilon \leq 0{,}037$ y $\eta \leq 0{,}058$).

Este corolario puede obtenerse fácilmente, a su vez, de un resultado debido a J. V. Uspensky, *Introduction to Mathematical Probability*, 1937, que da, en pp. 101-103, una extraordinariamente sencilla prueba de una versión fortalecida del teorema de Cantelli. Esta versión puede escribirse,

3) $$m \geq (2\varepsilon^2 - 4\log\varepsilon + 4\log 2 - 2\log\eta)/\varepsilon^2.$$

A partir de 3) obtenemos 2) bajo la condición de que las dos desigualdades $1/\epsilon \geq 4\epsilon 2 - 8\log\epsilon$ y $1/\eta \geq 8\log 2 - 4\log\eta$ sean satisfechas.

El teorema de Cantelli (a veces llamado, por ejemplo, por von Mises, «la ley fuerte de los grandes números», aunque Khinchine y Kolmogorov han descubierto desde entonces leyes más fuertes) nos permite, bajo la forma 3), calcular efectivamente para cualquier ε y η positivos, por pequeños que sean, un número *m* tal que la probabilidad de que (*) falle para cualquier *n* (mayor que *m*) será menor que η.

[22] *Probability, Statistics and Truth*, pp. 184-185. (2.ª edición alemana, pp. 154-157; 3.ª edición alemana, pp. 151-153.)

rema una interpretación que no es muy diferente de la interpretación frecuencial. Puede expresarse como sigue. [Me limito, en pro de la simplicidad, al caso de igual distribución $p\,(l) = 1/2$.]

Pongamos en una bolsa un espécimen de cada uno de los diversos segmentos de longitud n (donde n se elige de tal forma que $n > m$) de modo que haya 2^n segmentos en la bolsa, todos ellos diferentes. Entonces, entre esos segmentos habrá una fracción de, como máximo, 1/1.000 en la que las desviaciones de 1/2 excedan $\varepsilon = 1/1.000$ después del enésimo lugar, es decir, después del mil-millonésimo lugar de cada segmento.

Dada esta formulación del teorema, podemos entonces proceder a razonar como sigue:

Podemos escoger ε tan pequeña como queramos y si la hacemos cada vez más pequeña, entonces, como $m = 1/\varepsilon^3$, m llegará al infinito (y, por tanto, n también). Al final, es decir, con $\varepsilon \to 0$, nuestra bolsa no contendrá ya el conjunto de todos los posibles *segmentos de una longitud finita n,* sino *el conjunto de todas las sucesiones infinitas posibles* (por cierto, un conjunto no numerable, como lo muestra el «argumento diagonal» de Cantor). Al mismo tiempo, la probabilidad de encontrar una sucesión no convergente se habrá hecho cero.

Ahora bien, este argumento no puede reproducirse dentro de la teoría frecuencial en la que von Mises expresó nuestro teorema. En primer lugar, empezó con *una* sucesión, la alternativa *a* que cortó en número infinito de segmentos largos *n*. Pero uno no puede cortar *a* en infinitos segmentos de longitud infinita, ni siquiera en dos segmentos tales; cualquier división de *a* en segmentos, de los cuales al menos uno es infinito, puede producir, como máximo, *un* «segmento» infinito de *a*, que es *a* mismo (salvo un comienzo de segmento). Por tanto, no hay posibilidad de construir *b* de una manera como la que se usó antes; y fue sólo dentro de *b* donde von Mises definió la probabilidad de hallar (o no hallar) la desviación.

En segundo lugar, la idea de un colectivo *b* cuyos elementos son sucesiones infinitas no puede sostenerse dentro de la teoría frecuencial. En ésta, los elementos de un colectivo son, esencialmente, sucesos observables, o los resultados de experimentos. También puden ser sucesiones *finitas* de sucesos, porque una sucesión *finita* de sucesos puede interpretarse, a su vez, como un suceso complejo. Pero un elemento de un colectivo no puede ser, evidentemente, una sucesión infinita de sucesos.

Así, la transición efectiva al límite está totalmente bloquedada para la interpretación de von Mises. En la mayoría de los casos, esto no importaría en absoluto: el contenido de un teorema del límite puede, por regla general, expresarse totalmente sin hacer realmente la transición, hablando, simplemente, de

sucesiones finitas cada vez más largas. Incluso en el caso que tenemos entre manos, puede decirse lo mismo *si consideramos la interpretación clásica del teorema,* precisamente porque no nos permite proceder hasta el límite. Pero dentro de un marco que no nos permite proceder hasta el límite, el teorema no puede tener toda la fuerza que tiene dentro de un marco que hace posible la transición.

A esta crítica mía, un teórico frecuencial puede ofrecer la siguiente réplica: Admito, puede decir, que la interpretación de von Mises, aunque correcta en sí, no da toda la fuerza del teorema. Pero esto puede remediarse con una traducción más directa del teorema clásico al lenguaje de colectivos. Correspondiente a su bolsa de 2^n segmentos (con $n > m = 1/\varepsilon^3$ existe un colectivo *b*, uno de los segmentos escogido al azar en la bolsa. La probabilidad, es decir, el límite de la frecuencia relativa, de sacar un segmento de esta bolsa en el que ocurran desviaciones de 1/2 que excedan ε, será menor que ε, exactamente igual que en el modelo clásico. Además, esta probabilidad será cero si *m* y, por tanto, *n*, se hace infinito.

Hay que rechazar esta réplica, pero merece un análisis cuidadoso. Lo que muestra, correctamente, es que la probabilidad siempre irá vinculada a una idea suficientemente flexible de frecuencia: una idea que comprenda la frecuencia en clases finitas; su generalización, o sea, límites de la frecuencia en sucesiones infinitas; y ciertas otras generalizaciones, tales como las medidas definidas por conjuntos continuos. Pero, en la teoría frecuencial de von Mises, *sólo* se admiten los límites de la frecuencia en sucesiones infinitas de sucesos observables. En vista de ello, hay que decir lo siguiente:

a) Si admitimos el método, sugerido por mi hipotético oponente, de reinterpretar una frecuencia relativa en una clase finita construyendo el colectivo de saques al azar a partir de esa clase finita, entonces este tipo de reinterpretación se convierte en trivial y superflua. Porque no hace más que reenunciar un cociente finito ordinario en términos de una sucesión infinita de cocientes; y *además,* tiene que hacer la suposición de saques «al azar».

b) Nuestra bolsa original, y sus frecuencias, no tenían nada de fortuito. Este era un modelo puramente matemático con toda frecuencia calculable exactamente para todo *n*. No contribuye a la claridad el interpretar *esas mismas frecuencias* como resultados de saques fortuitos o aleatorios (es decir, independientes).

c) Dentro de una teoría frecuencial del tipo de la de von Mises, el enunciado de que los saques aleatorios darán esas frecuencias debe conservar su carácter de una *estimación hipotética*

sobre las frecuencias de sucesos fortuitos. Pero el enunciado correspondiente sobre el contenido de la bolsa lo constituían teoremas matemáticos demostrables.

d) Así pues, el enunciado de que *hay una probabilidad cero* de sacar de la bolsa de todas las sucesiones posibles de ceros y unos, y una sucesión con una frecuencia no convergente (un enunciado que difícilmente podría interpretarse dentro de la teoría de Von Mises) no es ciertamente equivalente al teorema en discusión, según el cual *casi todas* esas frecuencias convergen (siendo cero la medida del conjunto de las que no convergen).

Con lo que se ha dicho quedará claro que el teorema que hemos estado discutiendo —una forma de la «ley fuerte de los grandes números»— no equivale a una derivación del «axioma de convergencia» de von Mises. No establece (y no puede hacerlo) que *todas* las sucesiones converjan, sólo que *casi todas* convergen. El «axioma de convergencia» caracteriza a los «colectivos» como pertenecientes a la clase de sucesiones convergentes. En otras palabras, la teoría de von Mises singulariza, muy apropiadamente, una clase especialmente interesante de sucesiones.

La situación respecto al segundo axioma de von Mises, el axioma de aleatoriedad, es estrechamente análoga. Una vez más, no puede ser derivado tal como está. Pero puede ser reemplazado por el teorema de Doob[23], según el cual en el supuesto de que las sucesiones en cuestión se compongan de suceso *independientes*, todo sistema de juego falla en casi todas las sucesiones (en todas las sucesiones de sucesos independientes, con la excepción de un conjunto de medida cero).

El supuesto de independencia usado aquí es, en efecto, el mismo que mi supuesto de que el enésimo elemento es independiente de todos sus predecesores o insensible a la selección por los n-1 predecesores; y en este grado, resulta que mi intento de basar la teoría sólo en la suposición de independencia o de insensibilidad respecto a n (descartando el axioma de convergencia) estaba, en lo principal, en la línea acertada. Esto es importante en la medida en que la independencia o la insensibilidad respecto a n tiene un sentido intuitivo claro: tratamos de asegurarnos, sacudiendo el dado o mezclando o por métodos similares, de que el resultado de la enésima tirada o saque no estará influenciado, en ningún modo, por los resultados de las tiradas o saques anteriores.

[23] *Cf.* J. L. Doob, «Note on Probability», *Annals of Maths.* 37, 1936.

23. LA SIGNIFICACION DEL FRACASO

Tocamos aquí un punto crucial. Es aquí donde el fracaso de la teoría frecuencial resulta realmente significativo.

El problema que llamé en *L. I. C.*, sección 49, «*el problema fundamental de la teoría del azar*» es completamente resoluble en términos de la teoría neoclásica. Esta teoría, si se une a la interpretación propensivista, puede *explicar por qué* las sucesiones de sucesos independientes (sucesos que son insensibles a la selección por sus *n* predecesores) se comportan de la forma en que lo hacen: *por qué* se comportan como lo hacen los colectivos de von Mises; *por qué* sus frecuencias muestran una tendencia hacia la convergencia; *por qué* son al mismo tiempo aleatoriformes, de modo que (casi) todos los sistemas de juego fallan. El problema fundamental de la teoría del azar se resuelve cuando podemos entender este extraño comportamiento, esta regularidad *cum* irregularidad que, en cierto modo, esperamos todos de los sucesos fortuitos, pero que tiene que crear dificultades considerables a cualquiera que reflexione seriamente sobre ella.

La teoría neoclásica, combinada con la interpretación propensivista, muestra que podemos esperar ese comportamiento con una probabilidad igual a 1 en una sucesión infinita de sistemas independientes o fortuitos.

El problema se resuelve, fundamentalmente, siguiendo las líneas de Bernoulli, Poisson, Borel y otros teóricos clásicos; de hecho, siguiendo las líneas en las que yo traté de resolverlo en *L. I. C.* Porque en este sentido nunca estuve de acuerdo con von Mises. Mientras yo resaltaba el problema, poniendo quizá más enfásis del que nadie había puesto antes, von Mises creía que no había ningún problema y que teníamos que aceptar la existencia de sucesiones aleatorias como un hecho empírico definitivo. La teoría de la probabilidad, según su concepción, no hacía más que tomar nota de este hecho y describirlo en forma idealizada. Así fue como von Mises llegó a sus dos axiomas, que él consideraba irreducibles.

Se vio confirmado en su concepción cuando descubrió que los intentos clásicos de resolver el problema habían sido, todos ellos, circulares[24]. En parte porque deseperaba de romper esta circularidad y en parte porque, como positivista, no creía en la «explicación», renunció a la idea de una derivación: éstos eran hechos irreducibles de la naturaleza que podían describirse pero no explicarse.

[24] Véanse las secciones 48 (nota 6) y 62 de *L. I. C.*

Yo traté de reconstruir la teoría porque sentía, por un lado, que la crítica que hacía von Mises a la teoría clásica estaba justificada, pero que, por otro lado, *había supuesto más de lo que era necesario para la teoría matemática* y que con ello había hecho imposible la solución del «problema fundamental del azar». Yo intenté, pues, reducir sus supuestos al mínimo, tratando de mostrar que una teoría de las sucesiones fortuitas podía derivarse, esencialmente, de la idea de insensibilidad respecto a *n*, una idea equivalente a la idea clásica de independencia[25]. Pero ahora creo que las objeciones de von Mises contra la teoría clásica ya no son válidos, si consideramos la forma que ha adoptado la teoría, en parte gracias a las propias ideas de von Mises, en la obra de Cantelli, Kolmogorov, Wald, Church y J. L. Doob. Y, aún más, creo que con ayuda del teorema de la exclusión de los sistemas de juego, de Doob, el problema ha alcanzado, o casi alcanzado, una solución final (siempre que supongamos propensiones).

24. LA TEORIA NEOCLASICA Y LA FRECUENCIAL COMPARADAS

Como su predecesora clásica, puede decirse que la teoría neoclásica trata la probabilidad como una medida de posibilidades (o de propiedades, o de clases o conjuntos). Pero difiere de la teoría clásica en los puntos siguientes:

a) No parte de una definición de la probabilidad. En cambio, considera «probabilidad» todo lo que satisface las reglas de un cierto cálculo.

b) No necesita tratar a las distribuciones iguales o equiprobabilidades como más fundamentales que las distribuciones desiguales.

c) Sustituye, en un número de casos de importancia decisiva, ciertos *teoremas del límite* clásicos (que afirman que ciertas posibilidades *tienden* al límite cero o uno) por teoremas que afirman que ciertos conjuntos de sucesiones posibles tienen *las medidas* de cero o uno.

d) Permite otras interpretaciones, pero sugiere fuertemente, especialmente por su teoría de la independencia, una interpretación que *atribuye probabilidades a los sucesos y ocurrencias singulares* que se *contrastan* por medio de las frecuencias dentro de

[25] Esta es la razón de mi elección intuitiva de la insensibilidad respecto a *n* a la selección según los predecesores como distinta de la insensibilidad a la selección ordinal normal: la primera es una forma frecuencial de la idea de *independencia.*

sus sucesiones de repeticiones del suceso en cuestión, es decir, sugiere la interpretación propensivista de las probabilidades.

Trataré brevemente de estos puntos.

a) Se recordará (*L. I. C.*, sección 48) que la teoría clásica[26] *definía* la probabilidad como el cociente entre el número de casos favorables y el número de casos igualmente posibles. Viendo que «igualmente posibles» significa también aquí «igualmente probables», esta definición equivale a un intento de definir probabilidades no iguales en términos de probabilidades iguales; o, en otras palabras, equivale a la proposición de que el cálculo de probabilidades tome la equi-probabilidad como su concepto fundamental y construya el cálculo general sobre esos fundamentos. Esta sugerencia se discutirá más ampliamente en *b)*.

Por el contrario, la teoría neoclásica no intenta dar una definición de la «probabilidad» ni según las líneas de Laplace ni según las de von Mises. Separa claramente la tarea formal de construir un cálculo matemático de la probabilidad de la terea de interpretar este cálculo con vistas a sus aplicaciones habituales en los juegos de azar.

Tanto Laplace como von Mises, cuando definían la probabilidad tenían en mente las aplicaciones; Laplace tenía en mente las seis posibilidades de un dado y von Mises el extraño hecho de que, en largas sucesiones de tiradas, los lados de un dado caen irregularmente, pero con las mismas frecuencias. Para la teoría neoclásica, «probabilidad» significa cualquier cosa que satisfaga las reglas del sistema matemático formal. Por tanto, primero debe desarrollarse el sistema matemático (pensando tanto en la generalidad matemática como en las posibles aplicaciones); y la cuestión de sus diversas interpretaciones debe plantearse sólo después.

b) La teoría neoclásica no supone que la equi-probabilidad tenga un status preferente; y sobre todo, no intenta construir todas las probabilidades con sumas de «unidades» o «pedacitos» elementales equi-probables o, en otras palabras, como los resultados de *contar probabilidades.*

Como la concepción contraria sigue siendo muy popular[27] —uno oye con frecuencia que no existe otro método de evaluar el valor numérico efectivo de una probabilidad— puede merecer la pena añadir aquí algunos otros comentarios críticos. Pero antes de hacerlo, quiero recordar mi admisión (*cf. L. I. C.*, sección 57) respecto a que las estimaciones hipotéricas de probabi-

[26] La definición clásica suele atribuirse a Laplace, pero fue anticipada en esencia por De Moivre (1718).

[27] Véase, por ejemplo, L. Vietoris, en *Dialectica* 8, 1954, pp. 37 ss., esp. p. 43, nota 1.

lidades iguales, que pueden sugerirse por consideraciones de simetría, son de la mayor importancia física. (Estas, naturalmente, no son *ni derivadas ni derivables* de un principio de indiferencia; pero, como todas las hipótesis, pueden recomendarse a nuestra intuición por cualquier cosa —incluido el interesante hecho de que las equi-distribuciones maximizan la incertidumbre de las predicciones—.)

Mi primer punto es que (como a menudo destacó von Mises) incluso el simple hecho del dado cargado —evidentemente, un problema de importancia para la física— transciende el problema de la equi-probabilidad.

En segundo lugar, se recordará que muchos de los escritores clásicos que trataron sobre probabilidades (incluido el propio Laplace) y partieron de la definición de De Moivre, más tarde (algunas veces sólo unas páginas más tarde) empezaron a desarrollar una teoría más general totalmente independiente de la suposición de probabilidades iguales; derivando, por ejemplo, la fórmula binomial inmediatamente en su forma general.

Como tercer punto puede mencionarse la crítica de Janina Hosiasson[28]. Puede formularse considerando la diferencia entre dos juegos de azar muy similares.

Descripción del primer juego:

Tenemos una bolsa y dos urnas, I y II. En la bolsa hay *tres contadores*, dos de ellos marcados «I» y uno marcado «II». En cada una de las urnas I y II hay tres bolas; en la urna I, dos bolas son blancas y una negra; en la urna II, una bola es blanca y dos negras.

Sacamos un contador al azar de la bolsa. Si está marcado «I», sacamos a continuación una bola al azar, de la urna I; mientras que si está marcado «II», sacamos a continuación una bola de la urna II. El juego termina cuando hemos sacado una bola de cualquiera de las urnas, I o II. La cuestión es determinar la probabilidad de sacar una bola blanca.

La respuesta, naturalmente, es

$$(2/3 \cdot 2/3) + (1/3 \cdot 1/3) = 5/9$$

[28] Janina Hosiasson (nacida en 1899, apresada por la Gestapo en septiembre de 1941 y ejecutada en abril de 1942; su marido Adolf Lindenbaum, un distinguidísimo matemático y otros cuarenta distinguidos filósofos polacos sufrieron un destino similar) comunicó esta crítica a Sir Harol Jeffreys, que la recoge en su *Theory of Probability*, 139, p. 301. Hay una réplica a la crítica de Jeffreys de la teoría de contar las posibilidades favorables, de J. Neyman en su admirable *First Course on Probability and Statistics* (1950), pp. 21-24. Aunque la discusión y las soluciones de Neyman son, desde luego, perfectamente correctas, no ha comprendido, en mi opinión, el argumento de la Dra. Hosiasson y el Prof. Jeffreys. Por ello, voy a presentar de nuevo este argumento, en una forma ligeramente modificada y más elaborada.

Descripción del segundo juego:

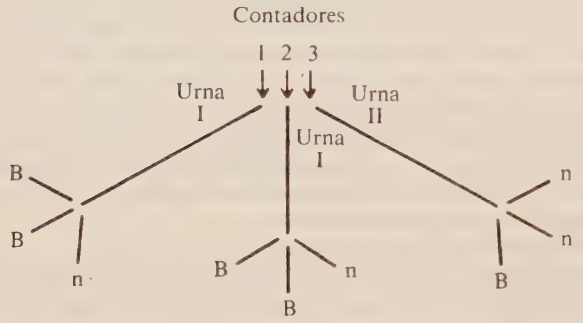
El segundo juego es exactamente igual que el primero, salvo que la urna II contiene sólo dos bolas, una blanca y una negra.

En el segundo juego, la respuesta a la cuestión de la probabilidad de sacar una bola blanca es, naturalmente,

$$(2/3 \cdot 2/3) + (1/3 \cdot 1/2) = 11/18.$$

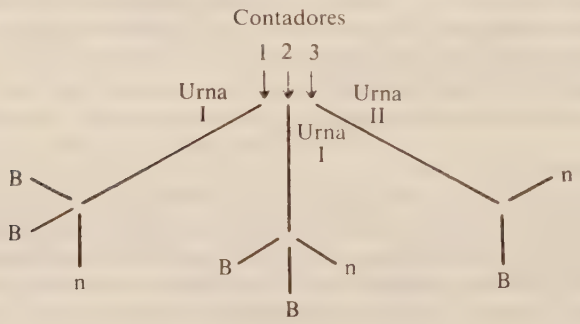
Comparando los dos juegos, encontramos que el primero puede representarse por el diagrama.

DIAGRAMA 1



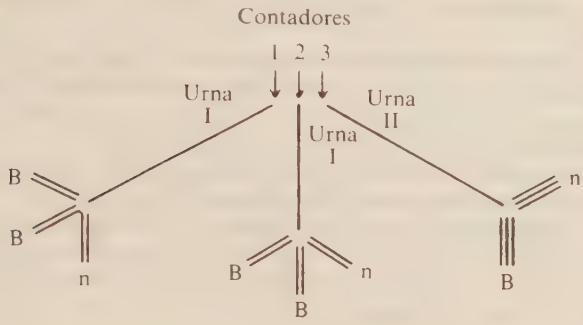
El segundo puede representarse con el diagrama

DIAGRAMA 2



Nuestro cálculo del resultado del segundo juego puede representarse por medio de un tercer diagrama,

DIAGRAMA 3



o, más bien, por la afirmación de que este diagrama es «*equivalente*» al segundo diagrama.

Ahora el argumento de Janina Hosiasson puede expresarse como sigue:

En el primer juego tenemos, esencialmente, nueve posibilidades iguales. (Como el diagrama claramente indica, podríamos no tener bolsas y en vez de una urna numerada con I, podríamos tener dos urnas diferentes, ambas marcadas con I, y, por tanto, tres urnas con nueve bolas en total, cinco de ellas, blancas.) Y podemos representar nuestro resultado diciendo que primero *contamos todas* las maneras iguales o posibilidades de terminar el juego —nueve en total— y luego *contamos* el número de las que terminan con el saque de una *bola blanca* (es decir, cinco). El cociente, 5/9, es la solución.

Pero este método es inaplicable al segundo caso. El resultado es 11/18. Sin embargo, es sencillamente falso que tengamos dieciocho posibilidades, de las cuales once sean favorables: ningún *recuento* de las posibilidades dará ese resultado. Aunque el segundo juego es muy sencillo (no hay la mínima dificultad para calcularlo) y aunque opera en cada etapa con iguales probabilidades o iguales posibilidades, su resultado no puede representarse dentro de una teoría que *cuente* posibilidades (iguales). Consiste, más bien, en una especie de *construcción* ficticia de dieciocho probabilidades iguales que podemos *calcular*, sin duda, pero que no se dan en absoluto en el segundo juego. Porque, aunque nuestro tercer diagrama describe un juego que, como todo el que sepa algo sobre el cálculo de probabilidades verá a primera vista, es «*equivalente*» al segundo juego, realmente lo que describe es un *tercer juego*, totalmente diferente del segundo; y la «*equivalencia*» del segundo juego (con ocho posibilidades desiguales) y el tercer juego (con dieciocho posibi-

lidades iguales) no es algo que pueda «verse» antes de la construcción del cálculo o que pueda «suponerse» básico para el cálculo.

El argumento puede resumirse así: En el segundo juego hay ocho formas diferentes de sacar bolas, cinco de las cuales producen una bola blanca. Pero las posibilidades se evalúan como desiguales (no importa la razón) y es el *cálculo* el que determina el resultado de 11/18 y no es recuento de las probabilidades supuestamente iguales. O, dicho de otro modo, aunque se formula el problema en términos de las posibilidades iguales existentes, su solución, a pesar de su simplicidad, no puede formularse de esta manera; lo que muestra lo inadecuado de cualquier definición basada en el *recuento de las posibilidades iguales*.

c) La sustitución, en la teoría neoclásica, de ciertos teoremas importantes del límite por teoremas que muestran que ciertos conjuntos tienen la medida cero (o la medida uno) es de considerable importancia filosófica. Por un lado, resuelve algunas de las supuestas paradojas de la teoría de la probabilidad y, por otro, ayuda a construir un «puente» que lleva desde las hipótesis probabilísticas a las contrastaciones en términos de frecuencias relativas.

Las supuestas paradojas han sido muy discutidas recientemente[29]. Están relacionadas con lo que se ha llamado en ocasiones el «principio de Cournot» (o el «lema de Cournot»); puede formularse como «los sucesos cuya probabilidad es muy pequeña son prácticamente imposibles»[30]. La dificultad relacionada con este principio (o con cualquier principio similar) surge del hecho de que los sucesos cuya probabilidad es muy pequeña ocurren en efecto, aunque raramente; o, en palabras de Paul Bernays, «uno puede, desde luego, presentar contraejemplos al lema de Cournot (tales como el que yo haya ganado la lotería)»[31]. (Además, toda ocurrencia efectiva suficientemente compleja tiene una probabilidad muy baja.)

Dentro de la teoría de la probabilidad, este problema se plantea en dos formas claramente distintas. Surge primero como el *problema del puente,* la transición de probabilidad a frecuencia; en segundo lugar, surge dentro de la teoría frecuencial (como problema de la contrastabilidad de los enunciados de probabilidades, tratado extensamente en *L. I. C.*).

---

[29] *Cf. Dialectica* 8, (núm. 30), 1954, pp. 125-144.

[30] *Cf.* Padrot Nolfi, *loc. cit.* p. 143.

[31] *Cf.* P. Bernays, *loc. cit.*, p. 140. De hecho, esta ocurrencia de un suceso arbitrariamente improbable (de probabilidad no cero) dentro de una sucesión de sucesos, puede hacerse tan probable como queramos haciendo la sucesión suficientemente larga; *cf.* por ejemplo, *L. I. C.*, sección 67.

Ahora, el primero de esos dos problemas puede resolverse completamente por medio de la nueva derivación de la teoría frecuencial a partir de la teoría neoclásica; es decir, por medio de la derivación de la ley fuerte de los grandes números y del teorema de Doob de la futilidad de los sistemas de juego como «casi seguros».

Porque éstos se derivan con una probabilidad *exactamente* igual a 1; de modo que las excepciones tienen una probabilidad *exactamente* igual a 0. Y no puede haber contra-ejemplos empíricos para un cuasi-principio de Cournot que se refiere a probabilidades *iguales* a 0: no hay billetes premiados con una probabilidad *igual* a cero.

Admitamos que esta falta de contra-ejemplos está relacionada con el hecho de que las nuevas derivaciones y las probabilidades cero se refieren a *sucesiones infinitas* y que no podemos producir una sucesión infinita de sucesos empíricos; tampoco podemos organizar una lotería con infinitos billetes. No obstante, las leyes de convergencia y de aleatoriedad para sucesiones infinitas constituyen la base de la teoría frecuencial; y para su derivación es muy importante evitar probabilidades que son meramente *próximas a* uno o *próximas a* cero.

Podemos ver esto con mucha claridad si recordamos la crítica de von Mises al «puente»[32]. Von Mises señala que, para utilizar los teoremas de Bernoulli o de Poisson como justificación para la deducción de enunciados frecuenciales a partir de enunciados de probabilidades clásicas sería necesario suponer *ad hoc* algún principio auxiliar tal como éste: «siempre que hayamos obtenido por nuestros cálculos una probabilidad que sea sólo un poco menor que 1, entonces el suceso ocurrirá en *casi todas* las repeticiones del experimento». Pero, señala von Mises, si hemos de interpretar una probabilidad de 0,999 como «casi siempre», ¿por qué no «conceder ya que una probabilidad de 0,50 significa que el suceso ocurre, por término medio, en 50 casos de cada 100»? O, en otras palabras, ¿por qué no adoptar ya la definición frecuencial?

Pero la nueva definición cambia las cosas aquí.

En primer lugar, elimina la siguiente dificultad: por ejemplo, no puede interpretarse satisfactoriamente que una probabilidad de 0,999 signifique «casi siempre». Por muy próxima a 1 que esté, se necesitaría algo parecido a un *límite* de la frecuen-

[32] *Cf. Probability, Statistics and Truth.* Cuarta Conferencia, la sección titulada «Supplementary Adoption of the Definition of Probability, etc.». Traducción inglesa (1939), y edición alemana, p. 135 s.; 3.ª edición alemana, p. 129. Véanse también las secciones 48 y 62 de *L. C. I.* (especialmente notas 6 y 3, respectivamente).

cia igual a 0,999 y también un enunciado de que esta probabilidad de 0,999 es insensible a los sistemas de juego. En otras palabras, necesitamos toda la teoría frecuencial si queremos interpretar 0,999 en el sentido frecuencial. El caso es diferente si obtenemos una probabilidad que sea *exactamente* igual a 1 (o a 0, como en caso de una medida cero). Hay que admitir que incluso en este caso, «probabilidad» tiene que significar algo relacionado con la frecuencia si queremos obtener el resultado requerido. Pero no hace falta asegurar una conexión precisa, ni un axioma del límite ni un axioma de aleatoriedad; porque se ha mostrado que éstos son válidos, excepto para casos que tienen una probabilidad (una medida) cero, y que, por tanto, pueden despreciarse. Así, todo lo que necesitamos suponer es que la probabilidad cero (o medida cero) significa, en el caso de los sucesos aleatorios, *una probabilidad que puede despreciarse como si fuera una imposibilidad.*

En segundo lugar, es bastante cierto que si interpretamos «probabilidad» de este modo, estamos obligados a conceder que una probabilidad de 0,50 supone, con probabilidad 1, que el suceso ocurre en 50 casos de cada 100. En realidad, el hecho de que esta conclusión está justificada es ahora demostrable. Pero ahora es demostrable sin implicar una definición de la frecuencia. Por tanto, no se plantea la cuestión de si la definición de frecuencia es *inadecuada.* Sencillamente, se ha hecho *innecesaria:* ahora podemos derivar consecuencias respecto a los límites de la frecuencia incluso sin suponer que la probabilidad significa un límite frecuencial; y hacemos posible así atribuir a la «probabilidad» un significado más amplio y más vago sin poner en peligro el puente sobre el que podemos movernos desde enunciados probabilísticos, de una parte, a enunciados frecuenciales que pueden ser sometidos a contrastaciones estadísticas, de otra.

Sin embargo, está claro que cualquier «puente» de este tipo, que lleve de la teoría clásica a la estadística, sólo puede construirse si la teoría clásica se interpreta *objetivamente* —por ejemplo, en el sentido de la interpretación propensivista—. A una interpretación subjetivista puede aplicársele mi antigua crítica (véanse sección 48, nota 6, y sección 62, nota 3, de *L. I. C.*).

d) Hemos visto que la teoría neoclásica *no define* la «probabilidad» y no intenta, por tanto, derivar el cálculo de una definición de «probabilidad» como hacían tanto la teoría clásica como von Mises. En cambio, construye primero el cálculo (bien de una forma axiomática o bien como parte de la teoría de la medida). Después pueden considerarse diversas interpretaciones del cálculo, sean subjetivas u objetivas. Pero el fracaso

de la interpretación subjetiva en la teoría de la independencia casi nos obliga a adoptar la interpretación propensivista.

No obstante, una interpretación frecuencial en el sentido preciso que sugirió von Mises (aunque, como ya dije, es una teoría coherente y sumamente satisfactoria) *no* proporciona una de las posibles interpretaciones de la teoría neoclásica: ésta última es auténticamente más general y contiene a la teoría frecuencial como una especie de «primera aproximación».

Como no parece haber otra interpretación objetiva fuera de las interpretaciones frecuencial y propensivista, y como la interpretación frecuencial no puede ser el «puente» hacia ella misma, no veo otra posibilidad que interpretar la teoría (neo)clásica en el sentido de la interpretación propensivista.

Respecto a esto, el aspecto principal es que la nueva teoría atribuye auténticamente probabilidades a *sucesos singulares*[33]. Y, aunque considera sucesiones de sucesos y frecuencias dentro de esas sucesiones, *la probabilidad de un suceso puede diferir radicalmente de su frecuencia en algunos de los segmentos observados de la sucesión*. (Sólo concuerda con *los límites* de la frecuencia *en casi todas* las sucesiones.)

He tratado de mostrar que una teoría del tipo de la de von Mises no puede ser considerada como una de las posibles interpretaciones del formalismo de la teoría neoclásica (fundada en la teoría de conjuntos) de la probabilidad[34]. Pero puede aparecer, a primera vista, que una reconciliación entre los dos enfoques —el neoclásico y el frecuencial— podría efectuarse. Por-

---

[33] Pienso que este punto no fue debidamente resaltado en las interesantísimas discusiones de 1940 entre von Mises y Doob; *cf. Ann. of. Math., Statistics* 12, 1941, pp. 191-217.

[34] Aunque, como acabamos de ver, hay resultados de la teoría neoclásica (conjunto-teórica) que no pueden interpretarse en una teoría frecuencial del tipo de la de von Mises, es posible interpretar un *sistema puramente formal* (en oposición a una interpretación conjunto-teórica de un sistema formal), tal como la expuesta en los Apéndices *II a *V de *L. I. C.*, en términos de una teoría frecuencial, por ejemplo, la del Apéndice *VI. Porque, sea S un conjunto de las más cortas sucesiones aleatorias (colectivos), tal que $a = a_1, a_2, \ldots$, $b = b_1, b_2, \ldots$, en donde cada elemento de las sucesiones, a_i o b_i, es igual o a 1 o a 0, y sea que S incluya dos alternativas compuestas de unos y de ceros sólo. Sea

$$p(a,b) = \lim ((\Sigma a_n b_n)/\Sigma b_n);$$

$$p(ab,c) = \lim ((\Sigma a_n b_n c_n)/\Sigma c_n);$$

$$p(\bar{a},b) = \lim ((\Sigma(1 - a_n)b_n)/\Sigma b_n);$$

$$p(a) = \lim ((\Sigma a_n)/n);$$

entonces todos los postulados y axiomas de *L. I. C.*, Apéndice *IV están satisfechos (y además de ellos, un postulado o definición de independencia).

que puede parecer que la siguiente *interpretación frecuencial* de la *teoría neoclásica* podría realizar la principal intención de un teórico frecuencial del tipo de von Mises:

(*) Interpretamos que la frase «la probabilidad (o medida) del suceso *x*» tal como se usa en la teoría neoclásica, significa «el límite de la frecuencia de los sucesos del tipo *x* dentro de casi todas las sucesiones de un conjunto infinito de sucesiones (aleatorias o aleatoriformes)».

Está claro que, dada una distribución de la probabilidad (o campo o espacio de calificación) la interpretación (*) puede llevarse a cabo. Esto se sigue, desde luego, de la ley fuerte de los grandes números y de la ley de Doob. Pero trataré de mostrar que (*) es totalmente insatisfactoria: que realmente significa empezar la casa por el tejado y que confunde de mala manera la situación.

La interpretación (*) no es, naturalmente, idéntica a la teoría frecuencial, pero es tan cercana a ella en su espíritu que un teórico frecuencial podría muy bien aceptarla como una especie de generalización de sus ideas: y podría argüir en favor de (*) como sigue:

«Probabilidad» significa, en primer lugar, frecuencia relativa dentro de una clase finita de sucesos, y en segundo lugar, el límite de las frecuencias relativas dentro de una sucesión infinita de sucesos. Con respecto a la sucesión infinita, podemos hablar de una «distribución», es decir, de las diversas probabilidades fundamentales cuya suma es 1, que no son más que los límites de la frecuencia de las propiedades fundamentales, o tipos de sucesos. Ahora bien, no hay ninguna razón por la que no podamos considerar, en lugar de una distribución que es relativa a *una* sucesión (como hemos hecho hasta ahora los teóricos frecuenciales, siguiendo a von Mises), una distribución que sea relativa a una *clase* de sucesiones, a saber, aquellas que tienen la misma distribución dada. Esto no cambia esencialmente la doctrina frecuencial, según la cual las probabilidades y las distribuciones sólo tienen sentido en relación con una sucesión de referencia dada. En esta nueva forma, tienen sentido en relación con *todas* aquellas sucesiones de referencia que pertenecen a un conjunto de sucesiones de referencia. Podría llamarse a este conjunto el «conjunto de referencia» de sucesiones de referencia de forma que incluyera «sucesiones excepcionales» con frecuencias diferentes de las probabilidades, siempre que la medida del conjunto de referencia fuese 1 y la del conjunto de sucesiones excepcionales fuese 0. De este modo, encontramos que (*) es una interpretación perfectamente natural del cálculo de probabilidades desde el punto de vista de la teoría frecuencial.

Pero este argumento confunde totalmente la cuestión.

Porque podemos hablar de probabilidades o de medidas en el sentido de la teoría neoclásica sólo en relación con una distribución fundamental, también llamada un «campo» o un «espacio»; y sin haber establecido esta distribución no podemos decir nada en absoluto sobre ninguna medida de un conjunto de sucesiones. Pero la distribución es una distribución de probabilidades. Empezamos, pues, con ciertas probabilidades cuando establecemos nuestro sistema de medición, que es un sistema de otras probabilidades. (Como siempre lo resaltó el propio von Mises, se nos dan probabilidades y de ellas derivamos otras.) El «conjunto de referencia», por tanto, tiene la medida uno y el «conjunto excepcional», la medida cero, *sólo* con nuestra distribución inicial, es decir, con las probabilidades que se nos dan al principio, y es empezar la casa por el tejado explicar esas probabilidades con respecto a un conjunto de sucesiones de medida uno que, para una distribución inicial diferente (es decir, dispuesta en un espacio diferente) se convertiría en un conjunto de medida cero. O, en otras palabras, si se nos da un conjunto continuo de todas las alternativas posibles, entonces, con respecto a una distribución inicial el subconjunto *A* puede tener la medida cero[35]. Es, por tanto, imposible utilizar (*) como una *explicación* (o definición) de «probabilidad» en términos de las frecuencias de *casi todas* las sucesiones; porque la «medida cero» traducida por «casi todo» resulta ser relativa a la distribución, es decir, a las *probabilidades* inicialmente supuestas. Así (*) ofrece una posible interpretación, sin duda; pero es insatisfactoria.

25. LA ESTRUCTURA DE LA TEORIA NEOCLASICA

Este punto es de cardinal importancia. Indica que una interpretación satisfactoria no será una que explique $p(x)$ en términos de frecuencias; o, dicho de otro modo, indica que *las*

[35] Si empezamos con $p(0) = p(1) = 1/2$, entonces obtenemos la medida uno para el conjunto A de alternativas con los límites de la frecuencia $p(0) = p(1) = 1/2$. Pero si empezamos con $p(0) = p(1) = p(2) = 1/3$, entonces la medida de A se hace cero, mientras que el conjunto B de sucesiones con las tres propiedades fundamentales 0, 1 y 2 y los límites de la frecuencia $p(0) = p(1) = p(2) = 1/3$ tiene la medida uno. Y si decidimos no distinguir entre las propiedades 1 y 2, y simbolizarlas a ambas con 1, de tal forma que empecemos con $p(0) = 1/3$, $p(1) = 2/3$, entonces el conjunto A vuelve a tener la medida cero, mientras que otro conjunto A', de alternativas con límites de frecuencia $p(0) = 1/3$, $p(1) = 2/3$, tiene la medida 1; pero en nuestro primer caso, A' tiene la medida cero. Así pues, la transición a una nueva distribución equivale a una transformación que no conserva la medida.

probabilidades de sucesos singulares u ocurrencias, aunque vinculadas en cierto modo a las frecuencias, tienen consecuencias que no agota la interpretación frecuencial.

Esto queda confirmado por el modo en que las sucesiones y las probabilidades de los sucesos están conectadas en la teoría neoclásica. La probabilidad de que un cierto suceso —digamos, el emésimo suceso en una sucesión de sucesos— tenga la propiedad *P* se introduce del siguiente modo.

Como primer paso, consideramos un conjunto fundamental de propiedades exhaustivas y que se excluyen mutuamente, *P, P', P''...* que puede presentar el suceso y coordinamos con cada una de ellas (libremente o, si se prefiere, hipotéticamente) un número positivo menor que uno, de tal modo que la suma de todos esos números sea igual a 1.

Estos números libremente asignados (o estimados hipotéticamente) resultarán ser las probabilidades de las propiedades pertenecientes al conjunto fundamental. Puesto que pueden escogerse libremente (o estimarse hipotéticamente), siempre con la condición de que su suma sea igual a 1, tenemos aquí una generalización del antiguo enfoque clásico (según el cual todos los números eran iguales).

Los dos pasos siguientes están pensados para introducir la teoría de la independencia (probabilística o fortuita) o de las probabilidades de ocurrencia juntas. Están basados, esencialmente, en la consideración de *todas las sucesiones posibles* y la atribución de medidas a ellas.

Como no conocemos la sucesión efectiva, construimos todas las sucesiones posibles, escribiendo:

1) todas las posibles propiedades que puede presentar el primer suceso de la sucesión real;

2) todas las posibles combinaciones de propiedades que puedan presentar los dos primeros sucesos de la sucesión real;

3) todas las posibles combinaciones de propiedades que puedan presentar los tres primeros sucesos de la sucesión real; etcétera.

Podemos llamar a estas posibles combinaciones de propiedades *sucesiones posibles.*

Está claro que si hay más de una propiedad fundamental,

a) el número de sucesiones posibles aumentará más deprisa que la longitud de éstas, de hecho, tan deprisa, al menos, como 2^n;

b) en el emésimo lugar, cada propiedad ocurrirá, al menos, en 2^{m-1} diferentes sucesiones posibles.

Pero esto significa que si consideramos una sucesión infinita de sucesos y, en consecuencia, un conjunto de sucesiones posibles que sean, cada una de ellas, de longitud infinita, entonces

a) el conjunto *U* de todas las sucesiones posibles será un conjunto infinito (de la cardinalidad del continuo);

b) el conjunto *S(m,P)* de todas las sucesiones infinitas posibles cuyo emésimo elemento tenga la propiedad *P* será también un conjunto infinito (de la misma cardinalidad).

Pero en el sentido de la antigua teoría clásica, deberíamos interpretar el número del conjunto *S(m,P)* dividido entre el número del conjunto *U*, como la *probabilidad* de que el emésimo elemento tenga la propiedad *P*.

Haciendo uso del hecho de que esos números son infinitos, asociamos al conjunto *U* la medida uno y al conjunto *S(m,P)*, una medida igual al número que, en el primer paso, coordinamos con la propiedad *P*. Así, el cociente clásico de las posibilidades se convierte en igual a este último número que podemos reconocer ahora como la probabilidad de que el emésimo elemento tenga la propiedad *P*.

Este es el segundo paso. Asocia la misma probabilidad con los sucesos emésimo y enésimo, para todo *m* y *n*, y asegura así la insensibilidad a la selección por el lugar.

El tercer paso establece que la medida de las sucesiones cuyos sucesos elésimo, emésimo y enésimo... tengan las propiedades *P*, *P'*, *P''*... respectivamente, es el producto de las medidas de los conjuntos de 1) las sucesiones cuyos elésimos elementos tengan la propiedad *P*, 2) las sucesiones cuyo emésimo elemento tenga la propiedad *P'*; 3) las sucesiones cuyo enésimo elemento tenga la propiedad *P''*... etc., es decir que establece la regla del producto, de la independencia.

De esto se sigue matemáticamente, entre otros teoremas más importantes, que:

i) El conjunto de sucesiones en las que las frecuencias no convergen con las posibilidades, tiene la medida cero. (Por consiguiente, el conjunto de las sucesiones con frecuencias no convergentes tiene la medida cero.)

ii) El conjunto de sucesiones que son sensibles a la selección según cualquier sistema de juego dado, tiene la medida cero. (Esto es un corolario del teorema de Doob.)

Como interpretamos la medida cero como una probabilidad cero, pero no como una imposibilidad, no hemos excluido la posibilidad de sucesiones que contradigan la teoría de von Mises; pero hemos mostrado que la probabilidad de que demos con tales sucesiones es cero.

En esta teoría, las probabilidades son medidas generalizadas de posibilidades; pero se muestra, con ayuda de lo que es, esencialmente, el método de Bernoulli, que las sucesiones con frecuencias que se desvían de las distribuciones de las probabilidades son tan raras que su ocurrencia puede despreciarse.

Por tanto, la interpretación de las probabilidades como medidas de posibilidades está enraizada en la estructura misma de la teoría neoclásica.

En la teoría neoclásica, tanto la probabilidad como la independencia son ideas que son lógicamente anteriores al cálculo de frecuencias; y no pueden reducirse a frecuencias. Sin embargo, los enunciados neoclásicos sobre la probabilidad o la independencia nos permiten afirmar con probabilidad 1 los enunciados cruciales sobre frecuencias que nos son necesarios en todas las aplicaciones a la física.

26. ENUNCIADOS DE PROBABILIDADES SINGULARES

La diferencia más importante entre la teoría frecuencial y la teoría neoclásica está en la interpretación de los enunciados de probabilidades singulares. La interpretación frecuencial de los enunciados de probabilidades singulares se ha discutido bastante extensamente en la sección 73 de *L. I. C.* Equivale a la aserción de que el enunciado «la probabilidad de que en la próxima tirada salga cara es de 1/2» *significa* lo mismo que la hipótesis de que «la frecuencia relativa de las caras en una sucesión (sea finita o infinita) de tiradas con esta moneda es 1/2»; es decir, la oración sólo parece ser singular, pero debe interpretarse como una oración sobre una sucesión.

En contraposición, la concepción neoclásica interpreta los enunciados de probabilidades singulares como enunciados que atribuyen probabilidades a sucesos singulares o, más exactamente, a un suceso singular *y* a un conjunto de circunstancias bajo las cuales se supone que el suceso en cuestión ocurre o no ocurre.

Llamaré a una teoría de la probabilidad «*teoría del suceso singular*»[36] si atribuye probabilidades a sucesos singulares o a enunciados singulares directamente, en vez de hacerlo por medio de un rodeo a través de sucesiones o conjuntos.

Uno oye decir a menudo que si queremos una «teoría del suceso singular» tendrá que ser una teoría que interprete la probabilidad como el grado de creencia racional, es decir, una teoría o subjetiva o lógica; sobre todo si hay que evitar una definición clásica (laplaciana), basada en la probabilidad o en la equi-posibilidad, por ser demasiado estrecha[37]. Parece ser

[36] *Cf.* E. C. Kemble, *Am. Journ. of Physics* 10, pp. 6 ss., Kèmble habla de «la probabilidad de los sucesos singulares».

[37] Véase, por ejemplo, E. C. Kemble, *loc. cit.*, e I. J. Good, *Probability and the Weighing of Evidence*, 1950.

que no pocos de los que creen en grados de creencia razonable lo hacen porque quieren adoptar una «teoría del suceso singular» y una que no esté basada en la equi-probabilidad. Creen (por alguna razón no muy convincente) que sólo puede satisfacer este deseo adoptando una teoría subjetiva o una lógica.

Pero no hay razón alguna por la que las concepciones clásicas o neoclásicas, según las cuales la probabilidad es una evaluación (o medida) de las posibilidades, deban estar vinculadas inescapablemente a la equi-probabilidad. Por el contrario, parece claro que la equi-probabilidad sólo se utilizó como modo de establecer una evaluación o medida y que nunca jugó más que un papel secundario en el desarrollo matemático de la teoría.

Haré unas cuantas observaciones críticas más sobre las teorías lógica y subjetiva de la probabilidad en una sección posterior. Aquí sólo quiero aclarar que no hay razón alguna para creer que si descartamos el enfoque equi-probabilístico, nuestra elección queda limitada a la teoría frecuencial por un lado, y a la teoría subjetiva o a la teoría lógica por el otro. Trataré de explicar mi punto de vista discutiendo primero un caso de equi-probabilidad —un dado perfecto— y variándolo luego al cargar el dado.

En el caso del dado perfecto atribuimos iguales probabilidades a cada uno de los seis resultados posibles de la próxima tirada. Esto es decir algo sobre la próxima tirada —un suceso singular—. El problema es descubrir qué es lo que se dice exactamente sobre este *suceso singular* (aparte de afirmar que pertenece, en potencia, a una sucesión con una cierta distribución de la frecuencia).

Sugiero que lo que afirmamos sobre un suceso singular —la próxima tirada— puede ser analizado en los siguientes términos:

1) Decidimos de antemano interesarnos por el resultado de la tirada sólo en la medida en que se caracterice por que quede arriba uno de los seis lados. (Así, descuidaremos, por ejemplo, la cuestión de qué lado queda hacia el oeste.) En otras palabras, delimitamos de antemano los «posibles resultados» bajo consideración.

2) Las condiciones objetivas bajo las cuales ha de ocurrir el suceso (o realizarse el experimento) son tales que no podemos predecir el resultado. (Este punto tiene una importancia secundaria, ya que podemos interpretar la plena predictibilidad como una aserción hecha con probabilidad uno.)

3) Imaginamos que a cada uno de los resultados posibles y sus combinaciones lógicas (estoy pensando sobre todo en su disyunción o su suma lógica) puede atribuírseles un número que satisfaga los axiomas del cálculo de probabilidades, espe-

cialmente el teorema de la adición, de modo que el número en cuestión pueda interpretarse como una medida aditiva; así, el número uno se atribuirá a la disyunción (unión) de todos los resultados posibles considerados y el número cero a la conjunción (encuentro, intersección) de dos resultados excluyentes.

4) Se pretende que esos números sean medidas de las diversas posibilidades que quedan abiertas por las condiciones del suceso o experimento: *si esas condiciones son, objetivamente, simétricas* respecto a esos resultados, pueden hallarse los números suponiendo que son iguales.

5) Es importante resaltar que en este análisis atribuimos iguales probabilidades sobre la base de una simetría objetiva supuesta, de las condiciones. Si estamos equivocados en el caso de un dado determinado (y todo dado físico real será al menos ligeramente asimétrico), entonces, nuestra atribución de números iguales, es decir, de probabilidades iguales estará equivocada en el mismo grado. Pero suponemos que:

6) A unas desviaciones muy ligeras de la simetría y homogeneidad corresponden unas desviaciones muy ligeras de la equi-probabilidad.

Este es el análisis del caso de un dado que se supone (aproximadamente) homogéneo. Atribuye los números (medidas de la posibilidad o probabilidad) a las *circunstancias objetivas* de un suceso; y no a nuestro *conocimiento subjetivo* ni al (objetivamente racional) grado de creencia que nuestro conocimiento pueda autorizar.

Para mostrar la diferencia, supongamos que se nos muestra un dado y se nos pide que determinemos la distribución de la probabilidad de los resultados de las tiradas. Según mi análisis, la respuesta correcta es: «No lo sé. Lo único que puede decir es que, *si* el dado es aproximadamente homogéneo, las probabilidades deberían ser aproximadamente iguales».

Veamos ahora un dado cuya carga pueda ajustarse con un mecanismo para que se mueva desde el centro hacia el lado opuesto al que marca «1».

1) Diremos que, para una posición central de la carga, se cumplirá aproximadamente la equi-probabilidad; y que a un cambio muy ligero de la carga corresponderá una pequeña (y quizá despreciable) desviación de la equi-probabilidad; y nuestro conocimiento de la mecánica nos sugerirá que la desviación significará un incremento de la probabilidad (posibilidad) de que quede arriba el lado «1».

2) Si nos preguntan «¿A qué llama usted una pequeña diferencia?», la respuesta apropiada, en mi opinión, es «no lo sé». No creo que podamos calcular la desviación de la equi-probabilidad a partir de la excentricidad conocida (como

sugiere Weyl)[38], porque la distribución de la probabilidad es una característica, o propiedad, no del dado cargado sólo, sino de todas las condiciones pertinentes; y puede depender, en parte, por ejemplo, de la superficie sobre la que cae el dado, de si es de acero o goma, o un almohadón plano cubierto de terciopelo, o arena, o barro. (*Cf.* sección 21, *supra.*)

3) Pero, aunque no podamos aplicar consideraciones sobre la simetría al cálculo de la distribución de la probabilidad en el caso del dado cargado, sí que sabemos algo sobre ella. Por ejemplo sabemos que el cambio de la carga, ya sea «pequeño» o «grande», ha incrementado la probabilidad del «1»; y esto entraña que *hay* una distribución de la probabilidad, aunque no la conozcamos, ligada a las condiciones peculiares del suceso singular.

4) No existe ninguna dificultad respecto a la aserción de que un suceso físico se caracteriza por números que no conocemos y que, con frecuencia, no podemos calcular. En matemáticas y en física matemática obtenemos muy frecuentemente resultados importantes discutiendo el comportamiento de una función matemática cuyos valores numéricos no podemos especificar (o no podemos hacerlo con precisión).

5) La propia incertidumbre que se atribuye a la determinación más exacta de la probabilidad me parece a mí una clara indicación de la *objetividad* de la probabilidad. Tomemos, una vez más, el caso del dado cargado. Podemos estar muy exactamente informados sobre la excentricidad y otras condiciones pertinentes. Pero este *conocimiento* nuestro puede no ser suficiente para determinar la probabilidad con el grado de precisión que queremos lograr, aunque puede permitirnos afirmar, por ejemplo, que la distribución es desigual. Habrá una probabilidad objetiva, pero no la conocemos, o no la conocemos todavía. Pero podemos saber cómo medirla, aunque quizá no tengamos tiempo, por ejemplo, de llevar a cabo la larga sucesión de experimentos que satisfaría nuestro deseo de lograr una determinación precisa.

Sobre la base de la teoría subjetiva, el estado de nuestro conocimiento determina la probabilidad *exactamente* en cada momento. Puede haber graves dificultades para calcular efectivamente la probabilidad a partir del estado de nuestro conocimiento, pero es un número exacto para cada estado de conocimiento, puesto que es una medida del estado de nuestro conocimiento o de nuestra falta de él. Según esta concepción, no tiene sentido hablar de medir *la* probabilidad con repeticio-

[38] *Cf.* H. Weyl, *Philosophy of Mathematics and Natural Science*, 1949, p. 197.

nes de experimentos, es decir, por medio de la adquisición de mayor conocimiento; porque un mayor conocimiento, en general, *alterará* la probabilidad. (Véase el capítulo anterior.)

Según mi concepción, la repetición del experimento dejará la probabilidad inalterada; de modo que podemos utilizar la repetición para mejorar la precisión de nuestra estimación de la probabilidad objetiva.

27. OTRA CRITICA A LAS TEORIAS SUBJETIVAS Y LOGICAS

Antes de comenzar mi análisis de la probabilidad objetiva de los sucesos singulares, tengo algo más que decir a modo de crítica de las teorías subjetivas y lógicas, desde un punto de vista ligeramente diferente del expresado en el capítulo anterior.

Estoy dispuesto a admitir que uno puede, sin encontrar problemas, interpretar el cálculo de probabilidades como un cálculo de *grados de expectativa racional*, o algo por el estilo, y que uno puede hacerlo o bien adoptando una perspectiva más subjetiva o psicológica (con Ramsey, Good o Kemeny) o una perspectiva más lógica (con Jeffreys, Keynes o Carnap)[39].

Lo que voy a criticar aquí es la concepción de que estas interpretaciones describen el uso de la probabilidad en las ciencias físicas.

Desde mi punto de vista, los enunciados de la física son objetivos y no se refieren de ningún modo a nuestro estado de información: ni «expresan» nuestra información ni nuestra nescencia. Son aserciones —aserciones conjeturales, naturalmente— sobre el mundo.

Esto es igualmente válido respecto a los enunciados de probabilidades de las ciencias físicas. *No son el resultado de nuestra carencia de conocimiento.* La carencia de conocimiento no produce milagrosamente un conocimiento de las frecuencias, ni siquiera con ayuda de leyes sobre los grandes números.

Que las hipótesis sobre frecuencias —por ejemplo, sobre la intensidad de las rayas del espectro— son tan objetivas como cualquier otra hipótesis física, lo admitirán, quizá, algunos que, no obstante, rechazan la objetividad de las probabilidades de los sucesos singulares. Pero lo que yo sostengo es que la dife-

[39] Tengo un poco más de duda respecto a si la interpretación en términos de grados de *confianza* racional no será un asunto totalmente diferente; puede ser bastante correcto decir 1) que la medida de mi expectativa en una tirada de cara es de 1/2; 2) que la medida de mi confianza en el suceso es cero; y 3), que la medida de mi confianza en la evaluación (de la propensión de 1/2) es uno.

rencia es relativamente pequeña. Las probabilidades objetivas de sucesos singulares dan lugar (por la ley de Borel-Cantelli y por la ley de Doob) a enunciados objetivos de frecuencias. Por otro lado, las probabilidades que expresan un estado de creencia razonable sólo podrían dar lugar a enunciados sobre frecuencias razonablemente esperadas; «razonablemente esperadas» en un sentido lógico si la interpretación original era lógica o tautológica.

En el caso de las probabilidades objetivas de sucesos singulares, evaluamos primero las condiciones objetivas bajo las que ocurrirá el suceso singular; y luego, por una derivación matemática, las condiciones bajo las que se dan ciertas frecuencias. En el caso de las probabilidades lógicas y subjetivas, evaluamos primero la relación lógica entre los «datos», el enunciado de la información «dada» y el enunciado del suceso en cuestión, y luego, por medio de una derivación matemática[40], evaluamos la relación lógica entre esos mismos datos y ciertos enunciados de frecuencias. En consecuencia, lo que obtenemos continúa ligado a nuestra información; no derivamos, como lo hacemos a partir de la hipótesis de probabilidades objetivas, una *hipótesis sobre las frecuencias*, que haya que contrastar y, si es necesario, rechazar; pero obtenemos (siempre que el cálculo sea correcto) *un enunciado verdadero* (de hecho, uno tautológico) sobre el grado en el que nuestro conocimiento dado autoriza la expectativa de una cierta frecuencia: en lugar de una conjetura física obtenemos un truismo sobre el estado de nuestro conocimiento y lo que esto implica. (Véase el capítulo anterior.)

Consideremos de nuevo el caso de un dado. Las teorías subjetivas y lógicas no tratan en realidad sobre el problema que le interesa al físico: ¿cómo se comportará este dado? No plantean realmente el problema (admitamos que es insoluble): ¿qué lado saldrá en la próxima tirada? La pregunta que se hacen es una pregunta diferente; es ésta: «¿Hasta qué grado nos permiten nuestros dados emitir uno u otro enunciado sobre el comportamiento de este dado?». Y respecto al próximo suceso singular sólo pueden preguntar: «¿Hasta qué punto nos permite nuestro conocimiento (o nuestra nesciencia) predecir que en la próxima tirada saldrá un uno?»[41].

---

[40] La derivación es muy cuestionable en este caso, porque no puede creerse «razonablemente» que la condición de la independencia objetiva esté satisfecha, si se interpreta como no pertinencia subjetiva, como lo resaltó Keynes por primera vez (véase *supra*).

[41] Muy a menudo, pienso que la mayoría de las veces, no se percibe la diferencia entre esas dos maneras de hacer preguntas; y cuando se percibe (como le pasa Jeffreys, a Ramsey y, quizá, a Good) a menudo lleva a la con-

A mí me interesa el comportamiento del dado y lo mismo les pasa, creo, a la mayoría de los físicos. Quiero saber algo sobre él, y estoy dispuesto a proponer conjeturas; por ejemplo, que los lados saldrán aproximadamente igual de frecuentemente en una larga sucesión de tiradas *hechas bajo ciertas condiciones;* o que las condiciones establecen una cierta simetría con respecto a todos los lados.

En contraposición, los teóricos subjetivos y lógicos están interesados en cuánto saben ya sobre el dado y no en las conjeturas sobre su comportamiento. Encuentran, por ejemplo, que su nesciencia en relación con el resultado de la próxima tirada es simétrica respecto a los seis lados y que, *por tanto,* las seis probabilidades son iguales.

Empezamos ahora una serie de tiradas y obtenemos, por ejemplo, 3, 1, 5, 1, 2, 2, 3, 5. El teórico subjetivo y el lógico afirmarán que estos nuevos datos deben afectar a la probabilidad de la próxima tirada. Sea mucho o poco, las probabilidades cambiarán: las de 4 y 6 decrecerán (mucho o poco); las de 1 y 2 aumentarán, porque 4 y 6 no han salido ninguna vez, mientras que 1 y 2 han salido dos veces cada uno.

Supongamos ahora que el dado es homogéneo y que la sucesión de tiradas resulta ser un estupendo ejemplo de una sucesión normal. Entonces las probabilidades —en el sentido de las interpretaciones lógicas y subjetivas— cambiarán siempre un poco, pero con el tiempo llegarán a establecerse cada vez más cerca de una equi-distribución. El efecto total de la larga sucesión de tiradas y de los nuevos «datos» habrá sido nulo. Así, los datos habrán resultado «no ser pertinentes en conjunto»[42].

Desde el punto de vista de la teoría objetiva —tanto si es del tipo del suceso singular como si es del frecuencial— la situación parece muy diferente. Considerada desde este punto de vista, la sucesión de tiradas constituye una contrastación

---

cepción de que la ciencia no es más que un instrumento para la transformación de nuestros «datos».

[42] Keynes vio siempre muy claramente este problema y lo describió como el problema del *peso de la evidencia.* El problema fue atacado por unos cuantos partidarios de las teorías lógicas y subjetivas, pero sin ningún éxito. Está claro que es bastante poco razonable decir que el enunciado probabilístico que atribuye la probabilidad de 1/6 al resultado *h* de la tirada con un dado tras muchas observaciones, no está mejor fundado en nuestra experiencia que el enunciado original sobre el resultado *h* de la tirada, basado en la simetría o en la nesciencia. Pero en la teoría de Keynes (y también en la de Carnap), este tipo de fundamento en la experiencia no está expresado. La teoría de Good, por ejemplo (*Probability and the Weighing of Evidence,* 1950, pp. 62 ss.) atribuye al ejemplo considerado aquí una «*ganancia neta del peso de la evidencia*» (p. 64) o de «la cantidad de información» (p. 63) (o una ganancia en «plausibilidad») que es cero.

estadística de una hipótesis. Si hubiéramos sospechado, originalmente, que el dado estaba cargado, la contrastación habría refutado nuestra sospecha. Así pues, podemos afirmar que corrobora nuestra conjetura original de igual probabilidad.

Consideradas desde este punto de vista, las interpretaciones subjetivas y lógicas confunden dos cosas: los enunciados físicos sobre sistemas físicos objetivos y las estimaciones epistemológicas sobre el grado en el que esos enunciados están «fundados en nuestra experiencia».

No es sólo la creencia equivocada de que toda teoría del suceso singular tiene que ser o subjetiva o lógica lo que explica esos errores. Parece que el hecho indudable de que nueva «información» *puede* cambiar las probabilidades de varios modos ha llevado a algunos pensadores a creer que las probabilidades no pueden describir las propiedades objetivas de las condiciones experimentales: si estamos «informados» sobre el hecho de que la tirada resulta en un número par, entonces la probabilidad de que salga «2» crece de 1/6 a 1/3; en símbolos:

$$P\,(2) = 1/6$$

$$p\,(2,par) = 1/3$$

Esto es así, indudablemente. Pero al interpretar la palabra «*par*» de nuestra fórmula como «información», ya hemos adoptado una interpretación subjetiva o lógica. (Esto no significa que no podamos utilizar el término «información», incluso en el contexto de la teoría objetiva, siempre que no nos dejemos llevar por nuestra terminología a mezclar dos interpretaciones.)

Desde el punto de vista objetivo, «*p* (2)» y «*p* (2,*par*)» se refieren a dos disposiciones experimentales diferentes: la primera expresión se refiere a una disposición experimental en la que se tiene en cuenta toda tirada; la segunda, a una disposición experimental en la que decidimos ignorar una tirada si es impar. También podríamos expresarlo así: la «información» expresada por el símbolo «par» nos dice que estamos haciendo una *pregunta diferente*, que ya no estamos preguntando: «¿Cuál es la probabilidad de que salga 2?» sino: «¿Cuál es la probabilidad de que salga 2, si sólo tenemos en cuenta las tiradas que resultan en número pares?»[43].

[43] Esta me parece una respuesta completa a los argumentos y ejemplos presentados por Erwin Schrödinger en el primero de sus artículos, «The Foundation of the Theory of Probability», *Proc. Royal Academy of Ireland* 51, 1945, sección A, pp. 51-66 y 141-146; véanse esp. pp. 63-66. El hecho de que en los ejemplos de Schrödinger la nueva información cambie algunas probabilidades pero no otras y establezca así ciertas nuevas «dependencias» o «pertinencias» no afecta de ningún modo a nuestro argumento.

28. LA INTERPRETACION PROPENSIVISTA DE LA PROBABILIDAD DE LOS SUCESOS SINGULARES

En las dos secciones anteriores traté de mostrar que los argumentos habituales contra la posibilidad de una teoría objetiva del suceso singular carecen de fundamento, incluso aunque la teoría sobre el suceso singular no se construya sobre la equi-probabilidad.

Pero la idea de la probabilidad como medida de las posibilidades sigue siendo, lo admito, algo débil; en especial es obvio que esta fórmula no puede ser considerada como una definición: «posibilidad», todos lo recordamos, no es nada más que otra palabra que significa «probabilidad»; en cualquier caso no es (como «frecuencia») una palabra cuyo significado sea más claro que el de «probabilidad».

No es mi intención *definir* la «probabilidad», y menos aún «la probabilidad de los sucesos singulares»: no necesitamos una definición (puesto que tenemos un sistema axiomático), sino simplemente una interpretación; y trataré de dar una interpretación y de hacerla intuitivamente más aceptable que la frase «medida de posibilidades». Propongo interpretar la probabilidad objetiva de un suceso singular como una medida de una *propensión* objetiva, de la fuerza de la tendencia, inherente a la situación física especificada, a realizar el suceso, a hacer que ocurra.

No tengo la menor duda de que este anuncio de mis intenciones causará horror a muchos de mis amigos positivistas, que verán en él una prueba de mis propias... he estado a punto de escribir «tendencias o propensiones» metafísicas, pero no debe hacerlo; así que digamos, quizá, «disposición» metafísica.

Yo no creo en la magia de las palabras y no me importa usar el término «disposición» en vez de «propensión» o «tendencia». Pero quiero resaltar que, como la mayoría de las hipótesis explicativas, una hipótesis sobre la probabilidad objetiva es *transcendente* en el sentido de que va mucho más allá de lo que puede saberse sobre la fase de las observaciones. (*Cf.* el final de la sección 25 de *L. I. C.*) Esto es válido para las hipótesis frecuenciales objetivas; y es válido aún en mayor grado (porque su grado de universalidad es mayor) para las hipótesis sobre probabilidades objetivas de sucesos singulares, puesto que éstas pueden *explicar* las frecuencias.

Para dilucidar la interpretación objetiva de los sucesos singulares puede ser útil empezar por señalar que una propensión 1 significará que es cierto o, al menos casi cierto, que el suceso ocurrirá, y que una propensión 0 significará que es cierto o casi

cierto que el suceso no ocurrirá: en estos límites ambas interpretaciones objetivas concuerdan bastante bien. Pero una propensión p con $0 \neq p \neq 1$ significa, en primer lugar, que un suceso del tipo contemplado puede ocurrir o no, bajo las condiciones consideradas; y junto con $p > 1/2$, significa que las circunstancias especificadas son tales que hacen más verosímil su ocurrencia que su no ocurrencia.

Pero esto plantea inmediatamente la cuestión de si no es *nuestra falta de conocimiento de las circunstancias precisas,* más que las circunstancias mismas, lo que da lugar a probabilidades distintas de uno o cero. Si la respuesta a esta pregunta es «sí», entonces debemos renunciar a la teoría objetiva. Si la respuesta es «no», entonces parece que podemos aplicar la teoría sólo a los sucesos que sean indeterminados en el sentido de que ni siquiera el más completo conocimiento de las «circunstancias» hará predictible el resultado. Pero, en realidad, queremos aplicar la teoría a tiradas con peniques y a otros sucesos macrofísicos parecidos, de los que nadie cree que sean indeterministas en este sentido.

Esta es una objeción importante, y una que obliga a los defensores de la teoría objetiva a ir a la raíz del problema.

Tomemos una máquina que tire monedas: uno pone un penique en una ranura, aprieta un botón y el penique sale de la máquina plano, sobre un almohadón blando. Si he visto hacer esto a la máquina una o dos veces y se me pregunta si estoy dispuesto a apostar por caras, la respuesta razonable sería, creo: «no sé si esta máquina es de las que 'aleatorizan' su resultado (como lo haría una mesa de ruleta bien construida). Por lo que yo sé, podría estar construida de tal modo que apretando el botón de determinada forma (o por algún método similar) uno pudiera producir a gusto el resultado que desea». La cosa es muy diferente si se deja al penique, ante nuestros ojos, rodar por una superficie rugosa hasta que coge velocidad y luego caer en un cubilete. En el primer caso diremos que dudamos de si las condiciones objetivas del experimento son tales que aseguren una «aleatoriedad» cierta en las condiciones iniciales mecánicas, mientras que en el segundo caso tendremos plena confianza en que lo son.

En ambos casos hay una falta de conocimiento. En el primer caso, las condiciones objetivas desconocidas *pueden* ser fácilmente tales que los que las conozcan puedan hacer predicciones precisas. De ese modo, el primer experimento puede ser predicho, y tal vez incluso controlado, por los que saben más sobre él que yo.

En el segundo caso, la situación es diferente. Forma parte de las condiciones del experimento el que aquellas condiciones

iniciales que podrían usarse para predecir el resultado sean «aleatorizadas». Con esto quiero decir que están arregladas de tal manera (superficie rugosa, etc.) que la mayoría de nosotros conjeturaría que, en una larga sucesión de experimentos realizados bajo las mismas condiciones especificadas (incluida la «aleatorización») es probable que las condiciones iniciales mecánicas cambien de una manera aleatoria.

Es esta conjetura positiva sobre el hecho de que las condiciones especificadas del experimento aseguran la aleatoriedad de las condiciones iniciales la que constituye la base de nuestra hipótesis sobre la probabilidad objetiva: conjeturamos que nuestras especificaciones experimentales son tales que producen condiciones iniciales aleatorias con tal de que se repitan.

Pero ¿qué pasa si hay un hombre que mida y calcule rápidamente las condiciones inciales reales de la moneda que rueda y de la superficie rugosa y hace cada vez la predicción correcta justo antes de que la moneda caiga en el cubilete? Mi respuesta es que sus mediciones no están en mayor conflicto con nuestras estimaciones sobre las probabilidades objetivas de los sucesos singulares de lo que estarían con una interpretación frecuencial: las frecuencias conservan la estabilidad y la aleatoriedad tanto si conocemos el resultado de los experimentos con un poco de antelación como si no. *Y lo mismo ocurre con las propensiones de los experimentos singulares* o, más exactamente, con las condiciones de la situación experimental singular. Porque con «propensión» quiero decir exactamente la disposición (o como se quiera llamar) de la situación *a producir esas frecuencias*, si el experimento se repite lo bastante a menudo. Las propensiones son disposiciones a producir frecuencias: ésta es la interpretación sugerida por la teoría neoclásica. Pero «propensión» *no significa* «frecuencia», porque hay sucesos que se repiten demasiado raramente como para producir nada parecido a un buen segmento de una sucesión aleatoria (o una «frecuencia»); con todo, esos sucesos raros pueden muy bien tener una propensión.

El número 1/6, en el caso del dado, se interpreta, pues, como característico de la disposición experimental y continúa siendo una característica válida de esta disposición aunque pudiéramos lograr, por precognición o por un cálculo rápido, predecir cada uno de los resultados en una larga sucesión de tiradas. El número se le atribuye a la disposición experimental singular sobre la base de lo que sabemos y no sobre la base de nuestra carencia de conocimiento. Un conocimiento adicional no interfiere con la probabilidad o propensión que caracteriza a la disposición experimental.

Queda una interesante cuestión: ¿por qué creemos que pro-

cesos tales como el rodamiento de la moneda sobre una superficie rugosa aleatorizarán las condiciones iniciales? Me propongo volver sobre esta cuestión en el tomo II del *Post Scriptum*, secciones 29 y 30 (la cuchilla de Landé).

Aquí sólo quiero hacer dos observaciones. La primera es que al hablar de una propensión quiero sugerir una idea intuitiva afín a la de una fuerza newtoniana, pero diferente de una fuerza en que produce frecuencias en lugar de aceleraciones. Las frecuencias cambian cuando cambian las propensiones. Las propensiones son tan «transcendentes» o «metafísicas» como las fuerzas newtonianas (a las que Berkeley acusó de «ocultas»). Matemáticamente son perfectamente definidas, son interpretaciones de un simple cálculo. Y en cuanto a su contrastabilidad, tenemos que apelar a los enunciados frecuenciales derivables de ellas (con probabilidad 1). Pero incluso el componente que transciende los enunciados frecuenciales es, en un sentido, contrastable, por medio de enunciados frecuenciales de un tipo diferente. (Esto puede mostrarse con un análisis de la teoría cuántica.)

Mi segundo y último punto es éste: La interpretación propensivista es, creo, la de la mecánica estadística clásica. Boltzmann habló de una *tendencia*. Pero creo que la interpretación propensivista describe mejor lo que los autores clásicos tenían en mente cuando hablaban del cociente entre los casos igualmente probables y los casos posibles: pensaban que este cociente era una medida (una medida especialmente importante y conveniente, aunque no la más general) de la propensión, característica de ciertas condiciones especificadas a producir un suceso dado.

Aunque la idea de la probabilidad lógica jugó un papel considerable en *L. I. C.*, admito que yo era un teórico frecuencial cuando escribí esa obra. Resalté que había múltiples interpretaciones del cálculo de probabilidades; pero continuaba creyendo que la interpretación frecuencial era de fundamental importancia para todas las aplicaciones en física.

Todavía podría seguir sosteniendo esta postura, pienso, en vista de la significación crucial de los teoremas frecuenciales derivables de la teoría neoclásica. Sin embargo, creo que quizá sea más correcto resaltar la continuidad que la discontinuidad de mis concepciones; porque he cambiado de opinión. Históricamente, el cambio ocurrió primero como consecuencia de mi continuado intento de entender la situación en la teoría cuántica (más sobre esto en los tomos II y III del *Post Scriptum*): fue ahí donde me di cuenta por primera vez de la necesidad de una interpretación propensivista. De ahí volví a la teoría de la probabilidad y descubrí, para gran satisfacción mía, que la teo-

ría neoclásica había proporcionado, en efecto, la base matemática para una interpretación propensivista con la construcción de ese «puente» entre las teorías clásica y frecuencial que yo había considerado antes, bajo la influencia de von Mises, imposible de construir.

En efecto, la introducción de la interpretación propensivista desmuestra distar mucho de ser *ad hoc*. Resuelve el «problema fundamental de la teoría del azar» (*cf. L. I. C.*, secciones 49 y 64). Es decir, explica *por qué* casi todas las sucesiones finitas de tiradas con peniques se comportan de la sorprendente manera en que lo hacen: por qué las sucesiones largas dan la sensación de que sus sucesiones relativas tienden a un límite; por qué combinan esta extraña regularidad con una irregularidad característica; y por qué sus segmentos parecen obedecer tan bien a la ley de Bernoulli.

Resumen final, 1982

1) El mensaje principal del último capítulo de este libro es que *el problema fundamental de la teoría del azar* (como lo llamé en *L. I. C.) ha sido resuelto ahora* de una forma mucho más satisfactoria que en *L. I. C.* Este perturbador problema puede expresarse:

¿Cómo podemos explicar que toda sucesión registrada de tiradas con una moneda o con un dado presente, por un lado, un carácter típicamente aleatorio y por otro, una frecuencia relativa estable que aparentemente tiende a un límite?

Este enigma puede resolverse, sugiero, si adoptamos la interpretación propensivista del cálculo de probabilidades. El propio cálculo resulta no ser más que una generalización métrica de la lógica proposicional (véase *L. I. C.*, nuevos apéndices *IV y *V) a la cual puede añadirse una conocida y satisfactoria definición de la independencia de los sucesos.

El cálculo, interpretado por medio de la interpretación propensivista, nos permite deducir:

A) Casi todas las sucesiones infinitas de sucesos independientes que ocurren bajo condiciones constantes tenderán a exhibir frecuencias relativas que tienden a límites que son iguales a las probabilidades (propensiones) de los sucesos singulares repetidos.

B) Casi todas esas sucesiones tendrán un carácter aleatorio descriptible de varias maneras, por ejemplo, por el fracaso, a largo plazo, de todos los sistemas de juego.

Estas dos características de las sucesiones de sucesos independientes A) y B) fueron *postuladas* en los años 1920 por el

matemático Richard von Mises[44], que los llamó axiomas y fundó en ellos una teoría matemática de la probabilidad. Pero ahora pueden ser deducidos, como he indicado, de un sistema simple, casi lógico, siempre que interpretemos las probabilidades como propensiones de los sucesos singulares a realizarse a largo plazo. Pero esto resulta ser una exigencia casi necesaria por otros argumentos.

Esta importante obra matemática pura fue iniciada por Jacob Bernoulli (*Ars Conjectandi*, 1713), que trabajó sobre ella durante veinte años, un logro espléndido. Fue continuada por muchos matemáticos. Yo la considero un éxito tremendo de la matemática pura y pienso que tiene el mayor interés filosófico posible. Esas ideas contribuyeron a resolver uno de los grandes enigmas del mundo, un enigma que Richard von Mises todavía consideraba insoluble.

2) También se muestra en este capítulo que este desarrollo es incompatible con la interpretación subjetivista del cálculo de probabilidades.

3) La concepción de que necesitamos recurrir a la probabilidad sólo a causa de nuestra carencia de conocimiento ha sido sostenida por muchos de los más grandes físicos desde Laplace; la había defendido Einstein en dos cartas que me escribió (una de ellas reimpresa en *L. I. C.*). Esta fue, sin duda, la razón de que yo recogiera en este libro tantos argumentos (repitiéndome a menudo, mucho me temo) en un intento de mostrar tan exhaustivamente como me fuera posible, que la interpretación objetiva resuelve el problema, mientras que la teoría subjetiva es incapaz de hacerlo.

4) La famosa objeción de Einstein al «dios jugando a los dados»[45] está basada, indudablemente, en su concepción de que la teoría de la probabilidad es un recurso temporal contra nuestra falta de conocimiento, contra nuestra falibilidad humana; en otras palabras, debida a su creencia en la interpretación subjetivista de la teoría de la probabilidad, una concepción que está claramente ligada al determinismo. Traté, en nuestro diálogo, de llamarle la atención sobre el hecho de que había que abandonar esta concepción; y según una carta de Pauli a Born, el 31 de marzo de 1954, Einstein (que ciertamente la sostenía al principio de nuestras reuniones) la abandonó[46].

---

[44] Difunto profesor de Matemáticas Aplicadas en Harvard.

[45] Carta a Max Born del 7 de noviembre de 1944, citada en Max Born, *Natural Philosophy of Cause and Chance* (Oxford, 1949), p. 122. Véase también la carta de Max Born del 9 de octubre de 1944, en *Albert Einstein, Max Born Briefwechsel 1916-1955* (Munich, 1969), pp. 207-208.

[46] Véase la discusión en *El universo abierto (O. U.)*, tomo II del *Post Scriptum* a *La lógica de la investigación científica*, pp. 26n-27n.

5) Puede añadirse que la solución al problema fundamental de la teoría del azar abre al menos una posible vía para la solución del problema de la causación. Que abre camino en aquellos casos donde (aparentes) leyes causales pueden ser derivadas como efectos-masa estadísticos (como por ejemplo, la ley de Boyle o las leyes causales basadas en la pequeñez del cuanto de Planck) es bastante obvio. Pero no parece haber ninguna razón para que no tratemos todos los efectos de carácter causal (todos los efectos deterministas *prima facie*, tales como el efecto de Compton) de la misma manera; es decir, tratando a las fuerzas como propensiones iguales a 1.

6) Naturalmente, he descubierto algunos nuevos argumentos desde que se concibió este volumen (aproximadamente en 1953). Algunos de ellos se encontrarán en los nuevos apéndices recientemente publicados en la séptima edición alemana, 1982, de *Logik der Forschung*. Espero incluirlos en la próxima edición inglesa de *The Logic of Scientific Discovery*.

INDICES

Compilados por Michael J. O'Donnell

Las referencias *en cursiva* indican los pasajes de especial importancia. El número de página seguido de *t* indica que el término se discute en el lugar referido; *n* significa «nota» y *c* (en el índice de nombres) significa «citado». Se aconseja al lector que consulte los índices de los demás tomos del *Post Scriptum*.

INDICE DE NOMBRES

Adams, J. C., 276, 277, 286.
Adams, W., 22.
Adler, A., 202, 209, 211.
Agassi, J., 21, 29*n*, 232*n*.
Albert, H., 187*n*, 208*c*.
Anaximandro, 99, 173*n*.
Anaxímedes, 173*n*.
Anderson, C., 28, 33.
Andersson, G., 41*n*, 108*n*.
Aristarco, 34*n*, 99.
Aristóteles, 30, 30*n*, 35, 45, 99, 176, 177*n*, 372.
Agustín, San, 117.
Avogadro, A., 230.

Bach, J. S., 123.
Bacon, F., 45, 54, 91, 121, 136, 139*c*, 146, 148*c*, 156, 231, 261, 273, 296, 372.
Bar-Hillel, Y., 380*n*.
Bartley, W. W., III, 22, 40*n*, 58 y *n*, 59, 60, 61, 67, 125 y *n*, 209.
Bayes, T., 42.
Becquerel, A. H., 32, 81.
Bellarmino, R., 156.
Bentley, R., 175*n*.
Berkeley, R., 118, 121, 124, 128, 129, 131*n*, 145 y *n*, 147*c*, 148*c* y *n*, 150*c*, 151, 152, 156, 157*c*, 158, 163, 165, 169, 172*n*, 220*c*, 221*c*, 231, 251*c*, 254, 304, 308*c*, 309*c* y *n*, 310*c*, 391*c*, 437*c*.
Bernays, P., 418*c* y *n*.
Bernfeld, S., 209.
Bernoulli, J., 341, 344, 346, 347, 354, 376, 397, 405, 406, 412, 419, 425, 439.
Bode, J. E., 98.
Bohn, H. G., 43*n*.
Bohr, N., 23, 28, 33, 36*n*, 166, 185*c*.
Boltzmann, L., 32, 200, 437.
Borel, E., 387*n*, 404 y *n*, 412, 431.
Born, M., 33, 180*n*, 187, 439.
Boscovich, R. G., 144.
Botticelli, S., 123.
Boyle, R., 87, 232 y *n*, 440.
Brahe, Tycho, 30, 99, 334.
Braithwaite, R. B., 322*n*.
Brewster, D., 297*n*.
Bridgman, P., 148*n*.
Broglie, L. de, 33, 232.
Brouwer, L. E. J., 110, 310, 312.
Bunge, M., 59*n*.
Burgess, T., 22.
Buridan, 30.
Butler, S., 125*n*, 212*n*.

Campbell, D. T., 138*n*, 262*n*.

Cantelli, F. P., 387*n*, 408*n*, 431.
Cantor G., 312, 409.
Carnap, R., 24*c*, 41, 149*nc*, 150 y *n*, 167*c* y *n*, 168 y *n*, 187, 219*c* y *n*, 235*n*, 252, 254*c*, 255*cn*, 268*cn*, 269, 292*cn*, 304*c*, 305*n*, 340*c*, 342*n*, 345*cn*, 347*n*, 355*cn*, 359*cn*, 361*cn*, 364, 368*cn* y *n*, 371, 372, 374*c*, 375*cn*, 377 y *cn*, 378*n*, 381*n*, 430, 432*n*.
Colodny, R. G., 249*n*.
Compton, A. H., 33, 440.
Comte, A., 191.
Copeland, A. H., 388 y *cn*, 392*n*, 400, 402*cn*, 404*n*.
Copérnico, N., 30, 34 y *n*, 36, 39, 148, 156, 230 y *n*.
Cotes, R., 175.
Courant, R., 307*n*.
Cournot, A., 418.
Craig, W., 160*n*.
Crooker, W., 81.

Challis, J., 277.
Church, A., 388 y *n*, 392*n*, 400, 402, 403 y *n*, 404*c*, 413.

Dahrendorf, R., 22.
Darwin, C., 39, 76, 64, 141.
Da Vinci, L., 315.
Davisson, C. J., 33.
Demócrito, 35, 45*c*, 99, 103, 231*c* y *cn*.
Descartes, R., 45, 99, 124, 142, 144, 175, 176, 232.
Dicke, R. H., 24.
Diels, H., 45*n*, 173*n*, 231*n*.
Dirac, P., 33.
Doob, J. L., 387*n*, 388*n*, 402*n*, 404*n*, 411 y *n*, 413, 421*n*, 422, 425, 431.
Duhem, P., 152, 158, 182*n*, 227, 228*n*, 231.

Eccles, J. C., 85*n*, 132*n*.
Ehrlich, P., 89.
Einstein, A., 24, 31-32, 33, 36*n*, 57*cn*, 65 y *n*, 66*n*, 78 y *n*, 79, 97, 99, 111, 115, 118, 120*c* y *n*, 131*n*, 152, 168, 171*c*, 185 y *n*, 189, 190, 193, 198, 200, 220, 232, 260, 270, 277, 287, 294, 303*n*, 371, 373, 381, 384, 385, 439.
Eötvös, R. von, 24.
Euclides, 243.

Faraday, M., 31, 144, 212, 232 y *n*.
Feigl, H., 149*n*, 187*n*.
Feller, W., 339*n*, 343*n*.
Feyerabend, P. K., 187*n*, 249*c*, 322*n*.
Findlay, A., 81*n*.
Fisher, R. A., 282*n*, 292.
Fizeau, A., 290.
Fleming, A., 81 y *n*, 88 y *n*, 89.
Florey, H., 89.
Foster, E. M., 58*c*, 62.
Franklin, A., 30*n*, 33.
Freed, J., 21.
Freed, L., 21.
Freeman, E., 46*n*.
Frege, G., 129, 169.
Fresnel, A. J., 31, 185 y *n*, 264, 271, 290, 294.
Freud, S., 202, 203*c*, 204*c* y *n*, 205*c* y *n*, 206*c* y *cn*, 207*c* y *cn*, 208*c* y *cn*, 209*c* y *cn*, 210*c*, 211 y *cn*, 212*c* y *n*, 213*c* y *n*, 304*cn*.
Frisch, O., 156.

Galileo, 30 y *n*, 31*n*, 34 y *n*, 39, 109, 110, 111, 148, 180 y *n*, 182, 183, 184, 188, 230 y *n*.
Galle, J. G., 277.
Geiger, H., 27.
Gerhardt, C. I., 265*n*.
Germer, L. H., 33.
Glasser, O., 81*n*.
Goldbach, C., 237, 238 y *n*.
Good, I. J., 426*n*, 430, 431*n*.
Goodman, N., 40, 108*cn*.
Gould, B. A., 277*n*.

Hadamrd, J., 75.
Hahn, O., 156.
Hanson, N. R., 33.
Harré, R., 31*n*, 34*n*.
Harrod, R., 321*n*.
Hegel, G. W. F., 78, 143, 195, 196, 216.
Heidegger, M., 217.
Heisenberg, W., 23, 166, 220*c*.
Hempel, C. G., 273*n*, 275, 296*cn*, 297*cn*.
Heráclito, 99, 78, 141, 143.
Hertz, H., 154, 290.
Hilbert, D., 250.
Hintikka, J., 269*n*.
Hobbes, T., 232.
Homero, 99, 211.
Hopf, E., 387*n*.
Hopkins, J., 347*n*.
Hosiasson-Lindenbaum, J., 288*n*, 415 y *n*, 417.
Hume, D., 52, 53, 61*c*, 62*c* y *n*, 71 y *n*, 72, 73, 74 y *n*, 75, 76, 77, 78, 79, 91, 92, 93, 94, 95, 96, 102, 112, 113,

115, 116, 117, 118, 119 y *n*, 120, 121, 122*c*, 128*c* y *n*, 129*c*, 131, 138, 140, 101*c*, 143, 144*c* y *cn*, 168, 169, 194, 195, 196, 259, 266 y *cn*, 304*cn*, 359*n*, 370*c* y *cn*, 386.
Husserl, E., 304*c*, 305*n*.
Huygens, C., 110.

Jaki, S. L., 287*n*.
James, R., 85*n*.
James, W., 126, 131 y *n*, 147.
Jeans, J., 23, 32.
Jeffrey, R. C., 262*n*, 381*n*.
Jeffreys, H., 53, 91*n*, 371, 372, 381*n*, 387*n*, 415*n*, 431*n*, 432*n*.
Jenófanes, 35, 51*c*.
Jesús de Nazaret, 124.
Jordan, P., 33.
Josué, 99.

Kant, I., 52, 53*c*, 77*n*, 78, 93, 94, 95, 118, 127*c*, 128, 191, 192, 193, 194, 195*c*, 196, 197, 295*c* y *cn*, 299*c* y *n*, 305*n*, 356, 371.
Keller, H., 139.
Kemble, E. C., 426*n*.
Kemeny, J., 269*c* y *n*, 362*n*, 430.
Kepler, J., 30 y *n*, 39, 62, 95, 98, 99, 148, 179, 180 y *n*, 182 y *n*, 183, 184, 186*n*, 187, 188, 230*c* y *n*, 334.
Keynes, J. M., 53, 107*n*, 267, 340, 371*n*, 430, 431*n*, 432*n*.
Khinchine, A. I., 408*n*.
Kneale, W. C., 216*n*, 392*cn*.
Kolmogorov, A., 408*n*, 413.
Körner, S., 153*n*, 322*n*, 392*n*.
Kraft, J., 186*n*.
Kramers, H. A., 33.
Kranz, W., 45*n*, 173*n*.
Kraus, K., 216*cn*.
Kruif, P. de, 89.
Kuhn, T. S., 35, 36*c* y *cn*, 37*c* y *cn*, 38*c*.

Lakatos, I., 26*n*, 108*n*, 237, 270*n*, 321*n*, 375*n*.
Langford, C., 305*n*.
Laplace, P., 380*n*, 414 y *n*, 439.
Lavoisier, A., 31.
Leeuwenhoek, A., 87.
Leibniz, G. W., 55*n*, 56, 118, 144, 265*c* y *n*, 307-10.
Lenard, P., 32-34.
Leucipo, 30, 144, 231*c*.
Leverrier, U. J. J., 276, 286.
Lewis, S., 89.
Lindenbaum, A., 415*n*.
Locke, J., 266.
Lorenz, H. A., 31, 32, 100.
Lummer, O., 32.

Mach, E., 32, 126, 131 y *n*, 145 y *n*, 147, 148*n*, 157, 158, 165, 172*n*, 200, 220, 231, 374.
Malthus, T., 76.
Marsden, E., 27.
Marx, K., 196, 202.
Maxwell, G., 187*n*.
Maxwell, J. C., 31, 32, 44*c* y *n*, 154, 185 y *n*, 200, 232, 264, 271, 289, 294.
Medawar, P., 90*n*.
Meiner, L., 156.
Meinong, A., 311.
Mellor, D. H., 322*n*.
Mendel, G., 39.
Menger, K., 312, 400*c* y *n*, 401*c* y *n*.
Metchnikov, I., 88.
Michelson, A., 32, 220.
Mill, J. S., 46, 91, 121, 158 y *n*, 191, 234*n*, 374*n*.
Miller, D., 22, 333*n*, 345*n*, 366*n*, 370*n*.
Mises, R. von, 236 y *n*, 244*n*, 387*n*, 388 y *n*, 400, 402, 403*c* y *n*, 406*c*, 408*cn*, 410, 411, 412*c*, 414, 415, 420*c*, 421 y *n*, 422, 423, 425, 438*c*, 439.
Misner, C. W., 66*n*.
Moivre, A. de, 414*n*, 415.
Moore, G. E., 304*c* y *n*, 311*n*, 312, 316.
Morgenstern, O., 401 y *n*.
Moser, S., 187*n*.
Mozart, W. A., 123.
Muirhead, J. H., 56*n*.
Musgrave, A. E., 22, 70*n*, 108*n*.

Nagel, E., 66*n*.
Needham, J., 27*n*.
Neurath, O., 105*n*, 167.
Newton, I., 30, 31*n*, 39, 62, 78, 95-101, 110, 115, 117*c* y *n*, 118, 147, 148, 150, 151, 154, 157, 158, 171, 175*c* y *n*, 176, 180 y *n*, 181 y *n*, 182*c* y *n*, 183, 184, 187-90, 193, 197 y *n*, 220, 229, 230, 260, 270, 271, 276, 286, 294, 303*n*, 307-309, 384, 385.
Neyman, J., 415*n*.
Niven, W. D., 44*n*.
Nolfi, P., 418*n*.

Okham, W., 374.
Oersted, H. C., 31, 44*c* y *n*, 81.
O'Rahilly, A., 66*n*.

Oresme, 30*n*.
Orwell, G., 315.
Osiander, A., 156.
Ostwald, W., 32.

Pagel, W., 27*n*.
Parménides, 30, 54*c*.
Parton, H. N., 268*n*.
Pasteur, L., 80.
Pauli, W., 439.
Pavlov, I., 85*n*.
Peirce, C. S., 322*cn*, 347*c* y *n*.
Perry, R. B., 125*n*.
Petersen, A. F., 22.
Planck, M., 28, 32, 440.
Platón, 35, 45, 54, 95, 99, 177 y *n*, 211*n*, 229, 303*c*, 304.
Poincaré, H., 81, 152, 158, 312.
Poisson, S., 412, 419.
Polanyi, M., 35.
Pólya, G., 387*n*.
Pötzl, O., 212*n*.
Pringsheim, E., 32.
Pryce, M., 321*n*.
Ptolomeo, 30, 34, 99, 101.

Radnitzky, G., 41*n*, 108*n*.
Rainich, G. Y., 66*n*.
Ramsey, F. P., 311*n*, 430, 431*n*.
Rayleigh, J., 23, 32.
Reichenbach, H., 121, 131, 144*c*, 163*c* y *n*, 164*c* y *n*, 165*c* y *n*, 166, 168, 371, 379*c*, 402*cn*.
Reid, T., 86*cn*.
Rembrandt, 123.
Robbins, H., 307*n*.
Roentgen, W. C., 32, 81 y *n*.
Russell, B., 22, 52, 53, 55 y *n*, 56 y *cn*, 71*n*, 93, 94*c* y *n*, 95*c* y *n*, 96, 102, 103*n*, 106*c* y *n*, 107, 120*n*, 121 y *n*, 126*c* y *n*, 127*c*, 128, 129*c* y *n*, 131, 163, 170 y *cn*, 215*n*, 298*c*, 305, 306*c* y *n*, 311 y *n*, 312, 371, 375*n*.

Schilpp, P. A., 37, 38*n*, 46*n*, 120*n*, 126*n*, 129*n*, 138*n*, 218*n*, 262*n*, 311*n*, 312, 378*n*.
Schlick, M., 152, 158 y *n*, 170, 234*cn*, 303*c* y *n*, 304 y *cn*, 401 y *n*.
Shönberg, A., 315.
Schrödinger, E., 33, 152, 232, 433*cn*.
Scott, R., 213*n*.
Scriven, M., 149*n*.
Selby-Bigge, L., 62*n*, 71*n*, 128*n*, 144*n*, 145*n*, 146*n*, 370*n*.
Settle, T., 22.
Shackleton, E., 213*n*.
Shearmur, J., 22.
Sheffer, H. M., 366*n*.
Simon, A., 33.
Slater, J. C., 33.
Sócrates, 54, 159, 215*n*, 254, 303, 304 y *n*, 332.
Soldner, J. von, 287*n*.
Soto, D. de, 30.
Spencer, H., 78*n*.
Spinoza, B., 253.
Sraffa, P., 371*n*.
Stekel, W., 211, 212.
Strachey, J., 204*n*, 206*n*, 208*n*.
Suppes, P., 66*n*.

Tarski, A., 23, 66*n*, 113, 119, 305, 307*n*, 313*c*, 314 y *n*, 366*n*.
Taylor, B., 306.
Thales, 173.
Thomson, G., 33.
Thomson, J. J., 27, 28, 31.
Thorne, K. S., 66*n*.
Torricelli, E., 30.

Urysohn, P., 312.
Uspensky, J. V., 408*n*.

Venn, J., 387*n*.
Vietoris, L., 414*n*.
Ville, J., 388*n*, 402*n*.
Vinogradov, I., 238, 243.

Waismann, F., 214*n*, 234*n*, 235*n*.
Wald, A., 388 y *n*, 392*n*, 400 y *n*, 401*c* y *n*, 402*c* y *n*, 403*n*, 413.
Warnock, G. J., 74*cn*.
Watkins, J. W. N., 21, 22, 228*cn*, 297*n*.
Weierstrass, K., 306.
Weyl, H., 31, 33, 429 y *n*.
Wheekler, J. A., 66*n*.
Whorf, B. L., 56 y *n*, 196.
Wien, W., 23, 30.
Wiener, N., 39*n*.
Wiener, P. P., 182*n*.
Wilde, O., 86*c* y *n*.
Wilson, E. A., 213*n*.
Wisdom, J. O., 204*n*, 205*n*, 213*cn*.
Wittgenstein, L., 152, 158 y *n*, 191*c* y *n*, 215*cn*, 234*c* y *cn*, 235 y *n*, 254*c*, 255*c*, 303, 304, 313.
Wrinch, D., 91*n*, 387*n*.

Young, T., 31, 110.
Yukawa, H., 28, 33.

Zenón, 358.

INDICE DE MATERIAS

A priori, 109, 111, 145 y *n*, 324.
validez, 118, 127, 176, 193-195, 356, 371.
y genética, 193.
Véase también: Apriorismo.
Absolución, regla de, 349, 353.
en lógica clásica, 349.
en teoría de la probabilidad, 349, 351, 378.
extendida a estimaciones estadísticas, 378.
Abstracto, abstracción y propiedades disposicionales de las cosas, 149.
Acción a distancia, 32, 184*n*.
Einstein y, 189.
Newton y, 189.
Aceleración, 180 y *n*, 437.
constante de, 182, 188.
de cuerpos de caída libre, 30*n*.
fórmula de Kepler para, 180.
teoría de, 30 y *n*.
Véase también: Cuerpos en caída libre, teoría de Galileo de.
Aceptabilidad, 282*t*, 322, 372.
de una hipótesis, 260, 293.
grado de, 273, 283, 372*n*.
Véase también: Corroboración, y confirmación.
Ad hoc, carácter-, en las hipótesis, 25, 95, 108, 109, 173-75, 184-85, 255, 262-72, 295, 419, 438.
Aleatoriedad, 387, 400, 403-04, 411, 420, 435-36.
Algebra booleana, 259*n*, 333.
Análisis, 316*n*, 317*n*.
de paradojas, 316*n*.
de significados, *véase:* Significado.
lógico, 235.
y lenguaje, *véase:* Lenguaje.
Angustia, 206, 207 y *n*.
síntomas de, 212.
Animismo, 304 y *n*.
Anomalía, Principio de Reichenbach de, 166.
Anticipación, *véase:* Expectativa.
Anticuerpos, 88*t*.
teorías de, historia de, 88-89.
Antropomorfismos, 396, 398.
Apoyo, explicaciones de, 108, 273-84, 365-66 y *n*, 367*t*.
Aprendizaje, Parte I, sección 3, 75, 79*t*, 80, 83, 138, 350, 354-56, 359.
por experiencia, 79, 86*t*, 140, 350, 354-57, 369-60, 362-64.
por imitación, un proceso de prueba y error, 83.
por inducción, 354-58, 359.
por prueba y error, 75, 79, 80-82, 85-86, 138.
por repetición o de memoria, 75, 79, 80, 82-85, 138-140.
tres tipos de, 80.
incapacidad inductivista a distinguir entre, 85.
y condicionamiento de los reflejos, 84.
Apriorismo.
de Kant, 127, 128, 193, 356, 371.
de Russell, 127, 128.
Véase también: A priori.
Aptitud, para la supervivencia, 109 y *tn*.
Apuestas, 350, 389, 392, 397.
y juegos de azar, 350, 351.
«Apuestas» de Reichenbach, 379.
Argumento, 46*t*.
circular, 173.
diagonal, *véase:* Argumento diagonal.
inductivo, 368.
justificador, y retrogresión infinita, 68.
método de, 203.
positivo o de apoyo, y refutación, 171.
racional, 67. *Véase también:* Argumento crítico.
no suficiente para mostrar la verdad, 66.
trascendental, 356, 359, 364, 369*n*, 372, 374, 379.
Véase también: Crítica.
Argumento crítico (o discusión crítica), 46, 60-68, 95, 99-102, 111, 117, 193-97, 215.

Argumento diagonal, de Wald y Cantor, 401, 409.
Aritmética, Frege sobre, 129, 169.
Arte, 315.
Artículos (publicaciones) sobre experimentos, 87-91.
Asociación, 72, 141, 158.
de ideas, 131.
ley de, 131, 147.
términos de, *véase:* Datos.
oposición del autor a las teorías de, 89.
Astrología, 30*n*.
cuestión de la evidencia inductiva en favor de, 109, 202.
historia de, 230.
pseudo-ciencia de, 230.
Atomismo, 199, 212, 215 y *n*.
como teoría metafísica, 231.
anterior a Demócrito, 212.
Atomo, atómos.
como vórtice, modelo del, 32.
propiedades estructurales del, 178.
teoría de Leucipo de, 30.
y el electrón de Thomson, 31.
y el vacío, 30, 231.
y moléculas, realidad de, 32.
Autoridad, recurso a, 61, 298, 299.
Axioma, axioma-esquema, sistema de axiomas, 160*n*, 258*n*, 259*n*, 421*n*.
Axioma del límite, de Ven y von Mises, 236 y *n*, 387*n*, 406.
discusión de Church, 403.
Véase también: Convergencia, axioma de (von Mises).
Axiomas de Hilbert, 250.
Azar, problema fundamental de la teoría del, 412, 438-40.

Bacteriología, 88.
Base empírica, 224-25.
Bernoulli, teorema de, leyes, métodos, resultados, 341, 344, 351, 376, 397, 419, 425.
como teoría subjetiva, 347.
Binomiales, leyes, 340, 345*n*, 415.
Biología, 232, 374.
estilo inductivo en, 87-89.
Bode, ley de, 98.
Bomba atómica, energía nuclear y, 299.
Borel-Cantelli, ley de, 431; *véase también:* Cantelli, teorema de.

Calculabilidad efectiva, 402, 403*n*.
Cálculo diferencial e integral, 136, 306, 308, 310, 316.
Cálculo de probabilidades, Parte II, 53, 59, 77, 187-88, 258*n*, 262, 268, 272, 277, 288, 291, 298, 322, 325, 331, 353, 364, 365-72, 372*n*, 380, 382-85, 427.
derivaciones de «probabilidad», 420.
interpretación clásica de, 244*n*, 326, 341, 396, 410, 413, 420, 426-27.
comparando a la interpretación propensivista, 396.
su aplicación a la física y a los juegos de azar, 335, 339, 340, 414.
von Mises sobre, 413.
interpretación frecuencial o estadística, 323, 324, 326, 327, 328, 333, 352, 377-78, 389, 390, 393-96, 399, 419-21 y *n*, 422, 427, 437.
abandono del autor de, 392*n*.
comparada con la interpretación de la medida teórica, 413-23, 426.
del autor, Wald y Church, 388, 392*n*.
donde fracasa y donde tiene éxito, 400-13.
historia de, 387-88, 399-406.
objeciones a, 391-95.
objetivas, 327, 334.
sustituida por la interpretación propensivista, 323, 327, 377, 387, 389.
y enunciados singulares, 327.
y von Mises, 408, 410, 411, 422.
interpretación de la medida teórica, 244*n*, 323, 387-88, 400, 411-14, 418, 420-21, 423-27, 436, 437.
carácter y comparaciones con la teoría frecuencial, 413-23, 426.
crítica del autor, 422-23.
interpretada en términos de teoría propensivista, 414, 420-21.
justifica la teoría frecuencial, 406.
interpretación lógica, 283, 323-24, 331-34, 337, 347, 365-66, 368, 372 y *n*, 377, 380, 427, 431, 432 y *n*, 433, 437.
como base de la interpretación subjetiva, 335, 336.
considera los enunciados de probabilidades como analíticos, 333, 334.
crítica del autor a, 430-33.
no se ocupa del comportamiento físico, 431-32.
interpretación objetiva, 321-40, 343, 347-48, 352-55, 370, 384-439.

comparada con la interpretación lógica, 335.
comparada con la interpretación subjetiva, 354-55.
depende de las condiciones físicas, 328, 353.
explica el funcionamiento de la regla inductiva simple, 347.
su tratucción a lenguaje subjetivo, 335.
y determinismo, 380.
interpretación propensivista, 322-23, 326*t*-28, 387, 389-93, 395, 399-400, 412, 421, 424*t*, 434-39.
como interpretación objetiva, 327, 335.
comparada con la interpretación clásica, 326-97.
comparada con la interpretación frecuencial, 323, 327, 399, 400.
interpreta físicamente las probabilidades, 326.
necesidad de, 437.
teoría neoclásica interpretada en términos de, 421.
y enunciados de probabilidades singulares, 327, 396.
interpretación subjetiva, 321-40, 341-42, 345-48,. 350-53, 358-59, 377, 393, 427, 429, 432 y *n*, 433.
basada en la interpretación lógica, 335.
comparada con la interpretación objetiva, 335, 341, 352-55, 357.
creencia de Einstein en, 439.
críticas del autor a, 260, 321, 421, 430-33.
interpreta la probabilidad y la inducción como generalización de la lógica deductiva, 380.
no distingue entre condiciones experimentales y el estado de nuestro conocimiento, 328, 335, 346-47, 393, 429.
no trata del comportamiento físico, 431-32.
núcleo de, 356.
su influencia entre los físicos, 341, 381.
y determinismo, 381.
Véase también: Probabilidad, probabilidades, teoría de la; Propensión, propensiones.
Cálculo inductivo, 368*t*.
carácter no tautológico de, 370-71.
coherencia de, 370 y *n*, 371.
Véase también: Probabilidad, explicación inductivista de la.
Cambio, 144.
Problema del, 199, 200.
Cantelli, teorema de, 404, 407*n*.
interpretaciones clásicas de, 408-09.
interpretación frecuencial de, 408-10.
no es derivación del axioma de convergencia de von Mises, 411.
Véase también: Ley de Borel-Cantelli.
Capacidad explicativa, 40, 66, 265.
Carácter satisfactorio.
Carencia de sentido (significado), 231, 248, 255.
Categórico, aparato categórico, crítica de Einstein, 57*n*.
Causalidad, 169.
crítica de Hume de la, 71*n*, 146.
física, 147.
postulado de Reichenbach de, 164*t*.
teoría metafísica de, 150.
Certeza, 53, 67, 142, 257-63, 367, 434.
absoluta, 225.
en la probabilidad, 328, 332, 335, 379.
lógica, 187, 379.
Ciencia, 96, 105, 110, 113, 141, 157, 170, 214, 254, 221, 294-301.
ataque a, 53, 54.
Berkeley sobre, 169.
carácter conjetural de, 109.
carácter revolucionario de, 36.
como conocimiento objetivo, 136.
como liberadora, 299-300.
como institución social, 136.
concepción verificacionista de, 224.
concepciones de Kuhn sobre, 35-39.
dudas del autor sobre la invariación de significados en, 249.
enfoque falsacionista de, 23, 224.
éxito de, 186.
explicación instrumentalista de, Parte I, secciones 11-14, 163-71, 299.
filosofía de, 169-70, 303, 374, 382.
fundamentos de, 306.
historia de, 28, 35, 38*n*, 179, 189, 199, 233, 316, 355.
Hume sobre, 169.
lenguaje de la, 255.
lógica de, 90, 201.
metafísica en, 121, 199, 219, 229-33.
método crítico de, 110, 161.
métodos de, 113.
objetivo de, 40, Parte I, sección 15, 73, 97, 156, 171-86, 282.
racionalidad de, 98.
realismo y, 186.

Schlick sobre, 303 y *n*.
su necesidad de teorías universales, 297.
teoría de, 27, 28-30, 151, 262.
teoría inductivista de, 29.
y naturaleza humana, 146.
y teoría del autor, 28, 171-72.
Véase también: Conocimiento científico; Método científico.
Ciencias físicas, 47, 400, 432.
Círculo de Viena, 253, 255*n*.
Claridad, valor de la, 47, 302, 315.
Colectivos, 236*n*, 388 y *n*, 401 y *n*, 402, 407, 408, 409, 410, 421*n*.
y prueba de Church, 388, 403.
y von Mises, 388*n*, 400, 402, 403, 411.
y Wald, 388*n*, 401.
como clase de sucesiones que convergen, 411.
insensibilidad respecto a los sistemas de juego, 402.
límite de frecuencia relativa de, 408.
Conceptos.
interpretación instrumentalista del autor de, 301.
interpretación esencialista de, 302.
ocultos, 148, *véase también:* Ocultos, teorías, efectos.
carácter metafísico de, 226, 254.
carácter operacional de, 149*n*.
científicos, 148, 253.
constitución de, 255.
versus enunciados, 149*n*, 215.
teóricos, 147-51.
universales, 150.
Condiciones, 358*t*, 431.
conocidas, 358.
generadoras, 395, 398, 399.
iniciales, 97, 114, 116, 174, 224, 236, 276, 291, 385, 436.
objetivas, 352, 353, 357, 358, 362, 389, 428, 435, 439.
pertinentes, 330, 331*t*, 359, 429.
reproducibilidad, 352.
Véase también: Condiciones, disposiciones experimentales. Independencia.
Condiciones, disposiciones experimentales, 87*n*, 329, 335, 337-40, 343, 348, 350, 357, 389-91, 398, 427-28.
condiciones generadoras de, 395, 399.
conocimiento de, 346.
elección de, 337.
propiedades disposicionales de, 391.
propiedades objetivas de, 433.
pertinentes para la contrastación de hipótesis, 328-31.
repetición de, 335, 337, 352.
variación de, 328-31.
y propensiones, 399, 436.
Véase también: Condiciones objetivas. Experimentos y repetición. Independencia.
Conducta, comportamiento y esencialismo, 176.
explicaciones de, 175-76, 431.
legiforme, 150, 178.
«plástica», 84*t*.
Conductismo, 24, 78.
Confirmación, 282, 298, 372.
grado de, 268, 292*n*, 322, 372 y *n*, 373.
y concepciones de Carnap, 167*n*, 292*n*.
y concepciones de Hempel, 273*n*, 296-97*n*.
y corroboración, 267-70, 284*n*.
y verificacionismo, 269, 274.
Véase también: Verificación.
Conjetura, conjeturas, 53, 126, 137.
comparadas con el conocimiento seguro, 126, 299, 381.
comparadas con el juego de azar inductivista, 381.
problema de la base subjetiva de, 127.
Véase también: Hipótesis.
Conjetura de Goldbach, 237*t*-48.
Conjetura de los números primos gemelos, 237*t*-49.
Conocimiento, 52*t*, 59, 85, 135*t*, 139--142, 201, 356, 359*t*, 367, 380, 429-30, 435.
adquirido, 137.
aumento del, 61, 67, 79, 82, 139-40, 195-96, 248, 358.
carácter disposicional de, 137, 140.
científico, *véase:* Conocimiento científico.
de sentido común, 132-33.
empírico, 52.
falible, carácter hipotético de, 93-94, 126-27, 141-42, 258, 260, 381.
fuente u origen de, 67, 140-41.
función de selección de, 38.
innato, 38, 134, 138.
insuficiente, 321, 335, 337, 341-42, 430, 435-37, 439.
objetivo, 119*n*, 134-36, 357.
personal, 133.
probable, 93.
rechazo del autor del, como creencia

verdadera justificada, 40, 114, 119*n*.
sin un sujeto conocedor, 132, 135.
sociología del, 24.
subjetivo, 81, 83, 134, 136, 140, 249, 328, 428-31.
teoría del, 51, 58, 62, 99-100, 114, 126, 128-29, 132-43, 151, 163, 165, 168-69, 171, 196.
valor del, 300.
y acumulación de evidencia, 293, 336-37, 340, 349, 351, 353, 357, 378, 393.
y experiencia, 111-12.
y leyes naturales, 111-12.
y teoría de la probabilidad, 328, 335, 341-42, 347, 393, 429.
y tradición, 137.
Conocimiento científico, 52*t*-53, 93, 121, 123, 127-43, 186, 263.
Conocimiento humano, 38, *passim*.
Conservación, leyes de, 30*n*, 33.
Conservación de la medida, 423*n*.
Contadores de posibilidades favorables, crítica de Hosiasson y Jeffrey, 415 y *n*, 417.
Contenido, 39 y ss., 174, 180, 256, 332, 366, 386.
empírico, 278, 285, 289.
lógico, 242-44 y *n*.
y probabilidad, 264, 266, 267, 271, 272, 278, 295, 323.
Véase también: Monotonía, ley de.
Contexto teórico, 218, 226.
Continuo, y números cardinales, 312.
teoría del, 310.
Contradicción, 285, 305, 309-10, 325.
Contrarios, teorías de Heráclito y Hegel sobre la identidad de los, 143.
Contrastabilidad, 40, 172-75, 214, 219, 235-51, 253, 275, 284-85, 295, 315, 418, 437.
criterio de, 51, 199, 214, 284*n*.
de teorías, 95-96, 229, 270, 295.
es igual a la corroborabilidad, 285, 289.
grado de, 180, 285.
versus sentido o significado, 214-18, 250, 253.
y carácter satisfactorio de las explicaciones, 185, 186.
y precisión, 315.
Contrastaciones, contrastar, 70, 73, 83, 161, 170, 173, 193, 227, 268, 275, 278-79, 283, 284, 287-90, 293-94, 296, 326-31, 348, 372, 375, 379, 381, 400, 420.
Convencionalismo 64, 192, 220, 315.
de Duhem y Poincaré, 152.
estratagemas del, 208, 315.
Convergencia, axioma de, (von Mises), 236 y *n*, 411, 419.
respuesta de Church a las críticas a, 403.
respuesta del autor a las críticas a, 405-06.
y posibilidad de sucesiones que lo contradigan, 425.
Correspondencia, principio de, 185.
y transición matemática de la teoría frecuencial a la de la medida teórica, y de la interpretación estadística a la propensivista, 387, 400.
de enunciado con hechos, 313.
Véase también: Verdad, teoría de la correspondencia de-, concepto de, Tarski sobre.
Corroborabilidad, 284*t*, 285, 288.
equivale a contrastabilidad y contenido empírico, 285, 289.
inversamente proporcional a la probabilidad lógica, 285.
contraria a la probabilidad, 289.
Corroboración, 51, 257-317, 372, 386.
como contribución a la metodología, 294.
como informe sobre la severidad de las contrastaciones, 290.
comparada con la probabilidad, 263--67, 271, 287-91, 295.
de hipótesis incoherentes, 291*n*.
debida a la emergencia de una teoría más fuerte, 271 y *n*.
definiciones de, 260-72 y *n*, 273, 279-94, 314.
desiderata del autor para, 282, 284--94.
grado de, 51, 98*t*, 109, 112, 260*t*, 263-67 y *n*, 270*t*-73, 283*t*, 294, 301, 322, 372*n*.
imposibilidad de análisis lógico completo, 275.
no lleva a alcanzar una meta, 300.
no satisface el cálculo de probabilidades, 267, 272, 282.
proporciona apoyo, 111, 279.
y confirmación, 267-70, 284*n*.
y equivalencia topológica, 282.
Cosmología, 99, 114, 118.
Cournot, principio de, 418.
Creencia, creencias, 61, 195, 355.
filosofía de, 62, 63.
filosóficas comunes, 170.
grado de, *véase:* Grado.

incredulidad del autor en, 61.
justificación de, *véase:* Justificación.
rivales, 59.
sistemas de, 197.
suspensión de, 119.
y acción práctica, 102.
y verdad, 119.
Véase también: Corroboración, grado de.
Creencia racional, 349.
explicación subjetivista del grado de, 293*n*, 336, 340, 349, 357, 427.
problemas de, 93, 96.
Cristianismo e idealismo de Berkeley, 124, 150.
Crítica, Parte 1, sección 2, 69.
apertura hacia, 67.
Bartley sobre, 67.
científica, 53, 67.
filosófica, 57.
problema de, 61.
solución del autor de, 61.
y problema de la justificación, 61.
forma en que Freud rechaza, 208.
inmanente, 69*t*, 70.
método de, 203*t*.
no justificacionista, 67.
normas de, 65, 92*t*, 97, 105, 311*n*.
nunca últimas, 69.
procedente de una teoría rival, 70.
racional, en la ciencia antigua y moderna, 35, 121, 220.
énfasis del autor en, 35, 67.
función de, 64.
ningún conocimiento exento de, 68.
y aumento del conocimiento, 67.
resistencia ante, no es razón para creer verdadera a una teoría, 60, 194, 270.
transcendente, 69*t*, 70, 380.
versus argumento positivo, 68-69.
Cuchilla de Landé, 437.
Cuerpo-mente, problema, teoría del monismo neutral del, 143.
Cuerpos (físicos).
caracterizados por leyes de la física, 131.
como procesos, 140.
esencia hipotética de 175.
existencia de, entraña existencia de regularidades físicas objetivas, 120.
idea de, 120, 129.
Realidad de, *véase:* Realidad.
rigidez e impenetrabilidad de, 178.
teoría de Galileo de la caída libre de los, 180-82, 188.
y mentes, reducción por Hume a manojos de ideas, 129.
Culpa colectiva, 212, 213*n*.

Choque cultural, 56, 194, 197.

Datos, 83-86, 141-42, 255, 432 y *n*.
Debate, *véase:* Argumento. Enfoque crítico. Argumento o discusión crítica. Crítica.
Definiciones, 251*t*, 306*t*, 420, 423, 434.
analítica, 316*n*.
cadena de, 254.
concepción del autor de, 39-40, 47, 372.
doctrina esencialista de, 175*n*, 176, 309, 315*n*.
elementales, 309, 311, 314*n*.
empíricas, 254*t*.
matemática, 248*n*.
regla de, 259*n*.
sistema de Carnap de, 254.
utilidad de, 23-30, 301-14.
y exactitud, 47, 302.
y precisión, 302, 314.
Véase también: Preguntas de qué-es.
Demarcación, entre ciencia y metafísica, entre ciencia y pseudo-ciencia, pre-ciencia, matemáticas, lógica, Parte I, capítulos II y III, 51, 214, 217.
criterio según la contrastabilidad, 23-27, 51, 149, 201, 214, 219-21, 229.
su aplicación a sistemas de teorías, 226.
no es criterio de sentido, 217-18, 235*n*, 255.
no es una distinción radical, 199, 201.
intento positivista de, 51, 149*n*.
problema de, 23*t*, 51, 95, 174, 199-203, 214, 216, 273.
distinto del problema de la verdad, 23.
y el desafío de Russell, 92.
Democracia, 300.
Derivada, definición de 305-09, 311.
Desatamiento, regla de, 349.
Véase también: Absolución, regla de. *Modus ponens*.
Descripciones, teoría de Russell sobre, 305, 311*t* y *n*.
Descubrimiento, 81, 89, 92.
Descubrimientos fortuitos de Roentgen y Becquerel, 32.
Determinismo, 119, 381, 400, 439.

Dimensión, dimensionalidad, 305, 312.
Disciplinas, 45-47, 200, 214.
Disposiciones, 396-98, 437.
activación de, 137.
físicas, 326, 390.
innatas, 137, 141.
metafísicas, 251.
programa de Carnap, para reducirlas a propiedades observables no disposicionales, 150.
Véase también: Expectativas. Conocimiento, disposicional. Propensión. Cálculo de probabilidades, interpretación propensivista.
Distribución, 288, 330, 422-25, 427-29.
Distribución estadística, 328, 330.
Divergencia lingüística, 197.
Dogmatismo, peligros de, 212.
de Freud, 208.
de Heisenberg, 220.

Efecto fotoeléctrico, 32-33.
Efectos secundarios, 329, 330, 344, 369.
Véase: Experimentos y efectos secundarios. Independencia.
Ejemplos, ejemplificación, 170, 203-04*n*, 208, 274*t* y *n*, 275, 296, 297*n*, 328, 368, 371, 376, 378-79, 419.
Electromagnetismo, 31, 33.
Electrónica, 31.
Elementos, supuesta inalterabilidad de, 33.
Elitismo, 35.
Empirismo, empirista, 118, 121 y *n*, 273.
clásico, 86, 139-40.
concepción del autor sobre, 79, 226.
ingenuo, 132.
principio de, 72*t*, 73.
Russell y Kant sobre los límites del, 52-53.
Enfoque crítico, 35, 38, 58-74, 75, 100, 110, 173, 196, 204, 273, 283, 300.
Entropía, interpretación absolutista de la ley de la, 32.
interpretación probabilística de, 32.
principio bayesiano de maximización de la, 385.
Enunciado, enunciados.
analítico, 285, 333, 348-50; *véase también:* Tautología.
atómico, 160, 366.
básicos, 24, 221, 225, 250.
carácter lógico o forma lógica de, 234*n*, 235.
condicional, 353.
contenido, 174.
contrastable, 237*t*, 241.
descriptivo, 152.
empírico, 254, 262.
estructura lógica de, 23.
existencial, 112, 114, 201, 218, 219, 221-23, 235, 236, 246-49, 272.
falsable, 23, 221.
frecuenciales, 400, 431, 437.
genuinos, 158.
infalsable, incontrastable, *véase:* Metafísica.
metafísico, 114, 218, 242, 243, 244.
observación de, 72-73, 262.
probabilidad, 283, 332-33, 334, 348, 382, 400, 418, 432*n*, Parte II.
proximidad lógica de, 332.
puramente existenciales, 24, 224, 233.
relaciones entre, y conceptos que ocurren, 222.
sentido, 217.
singulares, 167, 225, 246.
sintético, 350; *véase también:* A priori.
sustitución de universales por reglas de inferencia, 160.
universales, 159, 160 y *n*, 161, 217, 218, 221-23, 236, 247.
versus conceptos, 149*n*.
Véase también: Base empírica. Hipótesis.
Enunciados de probabilidades singulares, 426-30.
Epistēmē, 261.
versus doxa, 298.
Epistemología, 127.
carácter pre-darwinista de, 141.
de Kant, 193.
escéptica, 146.
inductivista, 168.
subjetivista, 121, 123, 124, 141, 144, 378*n*, 381.
Equidistribución, 333, 414-15.
Equiprobabilidades, 427-28.
como concepto fundamental de la interpretación clásica del cálculo de probabilidades, 413-14.
definición laplaciana clásica, 426.
desviaciones de, 428.
y desarrollo de la teoría neoclásica, 427.
Equivalencia, principio de Einstein de, 24.
Error.
aprendizaje por, 39.
detección de, 136.
sistemático, desconocido, 346.
Escepticismo, 58-61, 64, 120, 169.

Esencia, 176, 179, 301, 304 y *n*, 308.
Esencialismo, 174, 175*n*, 176, 177, 179.
de Freud, 304*n*.
metafísico y metodológico, 302*n*.
modificado, 178.
raíces animistas del, 304*n*.
Espacio, 220, 423.
Especialización, 300.
Espiritualismo, 131, 170.
Esquema conceptual, 36 y *n*.
Estadística, 346, 350, 353, 354, 360, 362, 420.
Estilo inductivo, Parte I, sección 3, 79, 87*t*-90.
Estímulo, 139, 180, 193.
Estructura, 170, 178, 179.
Evidencia, 275-83, 284 y *n*, 353-54, 357, 373, 380, 382-86.
empírica (u observacional), 40, 202, 224, 385.
explicación del autor de la, 293.
favorable, 40, 382-85.
idea de, y descubrimiento, 186.
independiente, 173, 185, 210*n*, 280.
inductiva, 100, 230, 258.
no afecta a las probabilidades relativas de las teorías universales, 385.
peso de la, 432.
Evolución, 194.
Exactitud, 217*n*, 298, 299, 429.
absoluta, inexistencia de, 310, 316.
no es un valor en sí misma, 47, 314.
y definiciones, 47, 302, 304*n*, 306, 312.
y resolución de problemas, 316.
Expectativa, expectativas, 81, 89.
estímulo de, 84, 137.
formadas por prueba y error, 84.
inconscientes, 32, 38, 83, 137, 193.
interpretación inductivista de, 83.
observación en, 85.
racional, grados de, 430.
Véase también: Suposiciones, conscientes e inconscientes.
Experiencia, 72, 86*n*, 87, 195, 214, 234*n*, 294, 351, 354-55, 371.
autoridad de la, 354.
como sistema de expectativas falibles, 87.
como supuesta fuente de conocimiento, 86*n*.
filosofía inductivista de, 87.
recurso a la, 71*n*, 214.
y observación, 75, 79.
y probabilidad, 342, 346, 378, 433.
Experimento crucial, 31, 34, 95, 106, 155.
Experimento de las dos ranuras de Young, 31, 399.
Experimento de Oersted, 31.
Experimentos, 90, 350, 352, 358, 378, 427, 429.
contrastación por medio de, *véase:* Contrastaciones, contrastar.
crucial, 400.
independientes, 329, 330-31.
singular, propensiones de, 436.
y efectos secundarios, 329, 330; *véase también:* Efectos secundarios.
y repetición, 329*t*-31, 337-39.
Véase también: Condiciones objetivas. Resultados experimentales; Repetición de un experimento.
Expertos, 45, 48, 299.
Explicación, hipótesis explicativas, 29, 224, 271-72, 423.
ad hoc, 173, 185.
causal, 172.
científica, 172.
circular, 173, 184.
de lo conocido por lo desconocido, 172, 231.
de regularidades antiguas y nuevas, 97.
definición como, 308.
no equivale a negar la realidad de lo explicado, 170.
estructural, 178.
explicans y *explicandum*, 172*t*, 173, 174, 184.
ideas y principios metafísicos, 30*n*.
problemas de, 40.
racional, 97.
satisfactoria, 109, 172-75, 186, 232.
teoría esencialista de, 176, 190.
últimas, 175 y *n*.
y contrastabilidad, 186.
y corrección, 185.
y leyes universales, 177.
y objetivo de la ciencia, 29, Parte I, sección 15.
Explicación (clarificación de significados).
abandono de, 305, 309, 310.
como objetivo de la filosofía analítica, 316.
insignificante en sí misma, 314.
no tiene un sentido preciso, 310.
tarea de, 304*n*.
Véase también: Definiciones. Explicación.
Extensión, 175, 178, 179*n*.

Falacia del jugador, 369*t*.
Véase también: Regla inductiva simple.
Falibilismo, y enfoque crítico, 38.
Falsabilidad, 23*t*-25, 26*t*, 36, 174, 213, 214, 221, 222.
carácter lógico de, 24, 25, 26.
criterio de, 174, 199, 213, 229.
dos significados de, 26.
leyendas históricas sobre, 28.
malentendidos sobre, 26-28.
no afectada por problemas de falsación empírica, 229.
no supone falsación en la práctica, 25.
se refiere a la posiblidad de falsación, 26.
y simetría, 223.
Véase también: Contrastabilidad.
Falsación, falsaciones, 25, 38*n*, 110, 209, 222, 230.
aceptación de, 23, 26, 29, 32, 34, 227.
carácter problemático de, 25.
contrastabilidad de, 27.
contrastaciones predictivas como, 372.
crean nuevas situaciones de los problemas, 33.
de éxito, 28.
descubrimiento de, 196.
dificultades prácticas de, 26.
ejemplos de, 26, 30-33.
estereotipo metodológico de, 37-39.
evasión de, 25, 220.
imposibilidad de, supuesta, 26, 27, 35, 38*n*.
incertidumbre de, 27.
llevan a reconstrucción de teorías, 29, 30-33.
negación el autor de las concluyentes o definitivas, 26.
no hay justificación final de, 39.
papel de, en la historia de la ciencia, 28.
potencial, problemas de su reconocimiento, 32.
problemas de, 26.
uso del término por el autor, 26.
y verificación, asimetría entre, 221-29, 262.
Falsacionismo, 35-39, 221.
estereotipo simplista de, 37.
ingenuo, 36-38 y *n*.
paradigma erróneo de Kuhn de las concepciones del autor, 37-38.
Falsador, y enunciados básicos, 26.
en conflicto lógico con teoría, 24.
existencia *versus* verdad de, 229.
potencial, 24*t*-27.
Fenomenalismo, 125, 200, 381.
Fenomenología, 304.
Fenómenos, 144, 148, 166, 180*n*, 254.
Fideismo, 36*t* y *n*, 61.
Filosofía, 171, 317*n*.
analítica, 316 y *n*.
contemporánea, 303.
crítica, 67.
choque del autor con doctrinas aceptadas de, 171,
empirista, 121-22, 128.
historicista, 177.
justificaciones, hasta ahora, 60, 67.
Kant sobre, 61.
natural, 146.
pre-kantiana, 196.
positivistas sobre, 254-56.
problemas de, 61, 202.
programas de Wittgenstein para, 234, 235 y *n*.
racional, 78.
Schlick sobre, 303.
tarea de, 58, 254, 304.
Física, 164, 327, 335, 339, 374, 381, 387, 397, 413, 427, 429.
ambigüedad en, 341.
carácter conjetural de, 430.
cuántica, *véase:* Física cuántica.
Descartes sobre, 175.
fenomenalista, 200.
filosofía de, 381.
historia de, 29.
leyes de, 112, 131, 165, 167.
meta-teoría de, 113.
nueva metafísica de, 400.
nuevo programa de investigación para, 400.
Reichenbach sobre, 163-65.
Física (mecánica) cuántica, 28, 32-33, 164, 166, 341, 400, 437.
Fisicalismo, 167.
Fisiología, 167.
Flogisto, teoría de, 31.
Forma.
lógica, y diferencia entre enunciados contrastables y enunciados metafísicos, 242 y *n*.
prioridad de las cosas individuales, 176.
teoría de Platón de, 177*n*.
Formación de hábitos, 82-83.
Formalismo, 166, 323, 376.
amor por el, 306.

como instrumento de predicción, 166.
de la teoría de la probabilidad, 403, 421.
de Lorenz, 32.
límites del, 220.
matemático, 242*n*.
Fórmula de indeterminación de Heisenberg, 220.
Fórmula de radiación, teoría de la, 23, 32, 185.
Frecuencia, frecuencias, 375, 388, 393, 395, 410, 411, 413-14, 418-21, 424, 425, 431, 434*t*, 436.
casi deducible de la probabilidad singular, 387.
estabilidad y aleatoriedad de, 436.
estadística, 390, 392, 397-98.
estimación de, 376-78.
explicación de, 434.
importancia en física, 426.
límite de, 419-23*n*.
observadas, observables, 347, 354, 391, 398.
propensiones contrastadas por medio de, 399.
relativas, 326, 342, 376, 389, 390, 395, 405, 410, 426.
como expresión de propensión, 326.
límites de, 236*n*, 392, 405-07, 410, 422, 438.
teoría objetiva de, 340.
y cálculo de probabilidades, 323.
virtuales, 396, 398.
y convergencia, 406, 411, 425.
y probabilidades subjetivas, 340.
Véase también: Cálculo de probabilidades, interpretación frecuencial o estadística de.
Fuerza, fuerzas.
campo de, 322*n*, 326, 391, 398.
como concepto relacional, 399.
concepto de, 184*n*, 391.
newtonianas, 31, 180*n*, 398, 437.
e hipótesis del autor, 309.
nucleares, y la teoría de Yukawa, 33.
vitales, 396, 398.

Grado de creencia, 336, 353, 428.
Véase también: Creencia racional.
Gravitación, gravedad, 31, 33.
y electromagnética, 31, 33.
explicaciones de Newton de, 109, 175, 189, 229, 270, 271, 384-85.
de Einstein, 31, 189, 270, 271, 384-85.
de Weyl, 31.
Grupos sociales, 195, 196.

Hechos.
conocimiento de, 163.
correspondencia con, 313, 386.
cuestiones de - *(quid facti) versus* cuestiones de validez *(quid juris)*, 76, 77.
dualismo de, 92.
empíricos, 193.
interpretados a la luz de las teorías, 180*n*, 193.
nuevas clases de, 157.
universales, 159.
Herencia, 39.
Hipótesis, 40-42, 62-63, 224, 228, 261-262, 352, 354, 372-75*n*, 382-84, 399.
aceptabilidad, *véase:* Aceptabilidad de una hipótesis.
análogo psicológico o biológico de, 137.
aptitud para superar las contrastaciones, 74*n*.
auténticas, 377-79.
auto-contradictorias, 285, 292*n*.
causales, e interpretación propensivista, 328-30.
contenido de, 262.
eliminación por medio de la crítica de, 109.
inductivas, 257.
matemáticas, 241.
método de, 372-74, 379.
probabilísticas, 328-29, 334. *Véase también:* Probabilidad, de las hipótesis.
pluralidad de, 109.
Véase también: Explicación, hipótesis explicativas.
Historia, interpretaciones de la, 211.
Humanidad, historia de, 47.
Humanismo, 294-301.
Hume, problema de, *véase:* Inducción, problema de la.

Idealismo, 120-26, Parte I, sección 8, 142, 150-51, 164-66, 171, 192, 197, 296, 304.
Ideologías, 57. *Véase:* Suposiciones mantenidas en común, 4.
Ilusiones, ópticas, 193-94.
Ilustración, 217.
Impenetrabilidad, 178 y *n*.
Improbabilidad, 284, 286, 367, 385, 418 y *n*.

lógica, 271, 287.
requisito de, 293.
y contenido, 270-72.
y simplicidad y capacidad explicativa, 265.
Incertidumbre, 257-63, 314, 429.
Véase también: Certeza.
Independencia, condición de independencia, 328-31, 336, 339, 340, 342-47 y *n*, 351, 364-66, 375-76, 410-13, 420-21*n*, 425.
Indeterminismo, 197*n*, 381, 435.
Indiferencia, principio de, 415.
Individuos, cosas individuales, 176-78.
Inducción, problema de, Parte I, capítulo I, 53, 62, 73, 74*n*, 93, 126, 143, 147, 158, 161*t*, 162, 180, 202, 257, 259, 270, 271, 361-63, 373*t*, 380-84.
autor sobre, Parte I, sección 2 y *passim*, 51, 53, 72-74, 76, 92, 102, 104, 116, 187, 199.
Bacon sobre, 300.
Berkeley sobre, 157.
Carnap sobre, 167, 187.
concepciones metafísicas impiden soluciones, 51.
creencia en, 74, 79, 298.
cuatro problemas de, relacionados, Parte I, secciones 4 a 6.
cuestión de su existencia, 52, 55, 115, 162.
Hume sobre, 52-53, 71*n*, 73, 75, 92--98, 102, 104, 140, 146, 158, 359*n*.
invalidez, imposibilidad de defensa racional de, 72, 75, 77, 111, 187--88.
justificación de, 74*n*, 91, 140.
mito de, 75, 161.
principio de, 52-53, 72*t*, 78, 115, 380.
regla simple de la, 232, 347 y *n*.
Russell sobre, 93-95, 103*n*, 107, 127, 371.
supuesto hecho de la biología y la psicología, 158.
teorías probabilísticas de, 167, 187, 259, 294, 341-86.
y deducción, 261.
y Hegel, 78.
y Kant, 93.
y Mach, 157.
y problema cuerpo-mente, 143.
y repetición, 82, 92, 191.
y verificación, 170.
Inductivismo, 57, 87, 91, 102-03, 109--10, 118, 180, 187, 262, 363, 372, 381.
Inercia, 30*n*, 175.
Inferencia, inductiva. *Véase:* Inducción.
Procedimientos inductivos:
licencias de inferencia deductiva, 113-14, 180, 257-58, 351.
Inferioridad, sentimientos de, 209.
Información, 431-34.
básica, 227 y *tn*, 275-84*t*, 292, 337, 340, 350-51, 357, 385.
Inobservables, inobservados, objetos, 163-64, 166, 391.
Insensibilidad, en probabilidades, 388, 401 y *n*, 402 y *n*, 403, 405, 412, 420, 426.
Instintos, como conocimiento innato, 134.
Instrumentalismo, Parte I, secciones 12-14, 175, 192, 231.
Interacciones electromagnéticas, 81.
Interfenómenos, 144*t*, 145*n*, 164*t*, 166.
Intersubjetividad, 127.
Intuicionismo, 54, 57, 310.
Irracionalismo, 27, 58, 61, 64, 67, 68, 391.
áreas de acuerdo del autor con el, 67, 68.
cómo puede uno verse arrastrado al, 140.
diferencias del autor con el, 60, 64, 67.
Irrefutabilidad, 122, 229.
Ismos, 54-55*t*; 300.

Juegos de azar, 335, 339, 340, 352, 372, 389, 415-18.
en oposición a apuesta racional, 350-51.
Justificación, justificacionismo, 39. Parte I, sección 2, 59, 62-64, 67-68, 91, 175, 336.
exigencia de, 69.
explicación de Bartley del, Parte I, sección 2.
imposibilidad de, 101.
las concepciones del autor erróneamente tomadas por, 61, 64.
mediante recurso a la inducción, 91.
no es problema de la teoría del conocimiento, 62.
positiva, 59, 63, 64, 69, 125.
problema de, 59-64, 111.
rechazo del autor de, 59, 62.
sustituido por el problema de la crítica no justificativa, 67.
y preferencias, 60, 111.

Lamarckismo, 135.
Lenguaje, 169, 197, 369*n*.

análisis del, 78, 145, 219, 235, 250.
artificial, 160, 314 y *n*.
carácter abstracto de los predicados, 149.
corriente, 254.
de relaciones, 56.
e idea del tiempo, 56.
incorpora teorías religiosas, metafísicas y empíricas, a menudo acríticamente, 55, 56, 59, 150.
influencia la filosofía, 56*n*.
meta-lenguajes y lenguajes objeto, 113*t*.
nos aprisiona, 56.
simbólico, 247*n*, 363.
subjetivo, 335.
universalista, 313*t*.
y la forma sujeto-predicado, 55, 56 y *n*.
y matemáticas, 242*n*.
y propiedades estructurales, relacionadas o disposicionales, 149.
Ley fuerte de los grandes números, 387, 404, 407-08 y *n*, 411, 419, 430.
Leyes de la naturaleza, 72, 103, 106, 113, 116, 118-19, Parte I, sección 6, 158, 161, 167, 169, 177, 186*n*, 191*t*, 192, 201, 234*n*, 360, 374, 438.
universales, 71-72, 74, 103, 106, 112, 114, 145-46, 158*n*, 168 y *n*, 174, 177, 192, 197, 236, 269, 292*n*, 296.
Liberación, proceso ilimitado de, 196--97.
Liberalismo, decadencia del, 196.
Lógica, 316.
deductiva, 162, 180, 259*n*, 309, 313, 333, 364, 365-66, 380.
inductiva, 53, 147, 258, 363-64, 365*t*-67, 371, 373-74, 377.
probabilística, 363-67*t*, 372, 375*n*.
Véase también: Cálculo de probabilidades, interpretaciones o teorías lógicas sobre.
Luz, cuanta de luz, (fotones), teoría de, 31, 33, 100, 110, 185 y *n*, 264, 271, 289-90, 294.
Véase también: Ondas, teoría de.

Magia, 252, 304.
Máquina en perpetuo movimiento, 221.
Máquina inductiva, 359-64.
Marco, marcos.
esquemático, para reacciones, 85.
herencia histórica de, 196.
instrumentalista, 157.
metafísico, 31*n*.
no irracionalismo, escepticismo, relativismo, 58.
susceptibilidad de ser mal interpretado el del propio autor, 58.
teórico, 195.
y vínculo con el relativismo, 195.
Marxismo, 24, 202, 204, 213.
Masa, 203.
Matemáticas, 237, 298, 306, 313, 316, 429, 439.
Materia, 144, 170, 200.
como última e inexplicable, doctrina de, 169.
historia de la idea de, 169.
Hume sobre, 143-51.
problema de, 143, 145.
propiedades esenciales de, 175.
realidad de, 143-51, 169, 170*n*.
estructura de, 144-45, 171.
teoría de campos de la, 169-70, 179 y *n*.
teoría de Mach-Oswald de, 32.
teoría eléctrica de, 33.
teoría electromagnética de, 28, 31, 212.
Materialismo, 131, 146, 169, 202, 212.
Mecánica clásica, interpretación estadística, propensivista y, 437.
Mecánica estadística, 341.
Medición, su limitado valor teórico, 184.
Mente, mentes, 169, 193-94.
como cubo, teoría de la, 139*t*, 140.
conocimiento de, 123.
Hume sobre, 144, 169.
naturaleza de, 129, 139, 304.
otras mentes, 123-24. *Véase también:* Solipsismo.
purgadas o abiertas, 54-55.
realidad de, 169-70.
Metafísica, 30*n*, 34*n*, Parte I, Capítulo I, secciones 4-8, 16, Capítulo II, 114, 120 y *n*, 214, 235*n*, 252, 295 y *n*, 391.
aristotélica, 304.
de Berkeley, 128.
de Hume, 128, Parte I, sección 8.
intento de eliminación de la, 219-20, 253.
miedo a la, 120.
observaciones lógicas sobre, 235-51.
platónica, 304.
programas de investigación de, 232*n*, 333.
y sentido (significado), cuestiones de, 215-16, 219-20, 254-56, Parte I, Capítulo III.

Metateoría, 63*t*, 64.
Método, metodología, 29, 34-39, 97, 101 y *n*, 109, 127, 171, 179, 201, 350-51, 361.
Método científico, 45-48, 94, 374, 380, *passim*.
Michelson, experimentos de, 31, 220.
Modas, el autor no cree en las, 47.
en filosofía, 55.
Modelo, modelos, 99, 364, 410.
Modelos atómicos.
de Bohr (1913), 23, 33.
de Rutherford, 27, 33.
de Thomson, 27-28.
Modus ponens, 349. *Véase también:* Absolución, regla de. Desatamiento, regla de.
Moléculas, estabilidad de, 191-92.
Monismo neutral, 125. Parte I, sección 8, 131*t* y *n*, 132, 143, 147, 163.
Monotonía, ley de, 259*n*, 264, 267. *véase:* Cálculo de probabilidades.
Movimiento, 30 y *n*, 105-06, 179-85, 188, 231, 334.
Movimiento browniano, 200, 329.
interpretación de Einstein del, 32, 200.
Movimiento planetario, 30, 32, 34*n*, 62, 180*n*, 182, 183, 334.
Multiplicación, ley general de la, 159*n*, 338, 384.

Naturaleza, 103, 145 y *n*, 146-47, 151, 169, 177-78*n*, 192-94.
Neptuno, descubrimiento de, 276, 277*n*, 286.
Nesciencia, *véase:* Conocimiento.
Neutrón, 33.
Newton, problemas de la mecánica y dinámica de, 99, 151, 180-84, 187-88, 190-92, 197 y *n*, 220, 276, 286.
No pertinencia, en la teoría de la probabilidad, 330, 331*t*, 336-40, 343-44, 431*n*.
Nominalismo, 302*tn*.
Núcleo, teoría del, 33, 156.
Número, números, 237-41, 424.
admisibles, 388*n*, 402*tn*; *véase también:* Sucesiones, normales.
enteros, 243*n*.
goldbachianos, 243, 246-48.
grandes, 350, 351, 407. *Véase también:* Ley fuerte de los grandes números, teorema de Cantelli.
naturales, 236*n*, 237, 238 y *n*, 239*n*, 240, 242-44, 245, 249.
primos, 237, 238 y *n*, 239 y *n*, 240.
teoremas de borel, 387*n*, 404*n*, 243, 249.
teoría de, 237-39*n*.

Objetividad, 88*t*.
Objetivista, objetivismo, 345-40, 350.
Observación, observaciones, 87, 95, 108-09, 164-65, 221, 250, 254, 326, 337, 381, 390, 432, 434.
acumulación de, 98, 138-39.
Carnap sobre, 150, 167.
como reacción compleja, 85.
descripción de, y estilo inductivo, 87.
e inducción, 53, 94-95, 121, 139, 158, 170.
en contrastaciones, 73.
evaluación de, 99.
falsadoras, 94-95.
fortuitas, 80, 81.
independencia de, 347.
interpretación de, a la luz de teorías, 36, 156, 204, 262, 273.
recurso a, 59, 214.
selectividad de, 100, 107.
significación sólo a la luz de teorías, 139.
sistemáticas, partiendo de un problema, 80.
y creencia racional, 107, 354.
y disposiciones y carácter disposicional, 139, 140, 150.
y evidencia, 113, 121, 132.
y repetición, 72, 75.
y subjetivismo, 133-34.
Ocultos, efectos, teorías, 148 y *n*, 157, 220, 251, 346, 391, 437.
Ockham, navaja de, 374.
Ondas, teoría de, 33, 154, 185 y *n*.
Ontología, en Hume, 129.
Organismos, 139, 398.

Palabras.
falta de interés del autor en las discusiones sobre, 63.
interpretación instrumentalista del autor de, 301-02.
significado de, 305*n*.
Paradigmas, 28, 29, 311*n*.
«Paradoja» de Goodman, 40-42.
Paradojas.
de Goodman, 40 y ss., 108.
de Hempel, 296*n*.
de la teoría de la probabilidad, 418.
del programa subjetivista, 380.
supuestas paradojas en Frege y Russell, 129.

Paranoia, Freud sobre, 209.
Partículas, 27, 33, 178, 192.
Penicilina, descubrimiento de la, 81, 88, 89.
Percepción, 131, 142, 145 y *n*, 147, 194; *véase también:* Observación.
Pertinencia, 336-39, 344, 358, 370 y *n*, 371.
Piedra filosofal, 233, 235.
Plasticidad, 85, 137, 140.
Platonismo, 304.
Poder explicativo, 291*t*; *véase también:* Capacidad explicativa.
Politeísmo, 229.
Posibilidades, 326, 327, 396, 426-28, 434-35*t*.
  iguales, 326, 396, 397, 417-18.
  interpretación física de, e interpretación propensivista de, 326, 396, 398.
  interpretaciones de, predictivas, 397.
  probabilidad como medida de las, 323, 326, 395.
  recuento de, 417.
Positivismo, 57, 125, 145, 149, 160, 216, 217, 219-21, 227, 235, 249, 253-56.
Positrón descubrimiento del, 33.
Potencialidades, aristotélicas, 398.
Pragmatismo, 64, 78.
Pre-ciencia, 23, 304.
Precisión, 355*n*.
  búsqueda moderna de, 366.
  descuido de, 310.
  exigencia de, 306.
  falso objetivo de, 47-48, 236.
  grado de, 174, 287, 355.
  no tiene significado preciso, 310.
  valor de, 185, 315.
  y definición, 302, 305.
  y contrastabilidad, 315.
Predicamento egocéntrico, y realismo e idealismo, 125 y *n*.
Predicados, 245-49, 251-54, 360-61.
Predicciones, 152, 155-56, 168, 284, 287, 328-29, 361, 368, 374, 382-83, 397, 415, 434-35.
Predictibilidad, 427, 435.
Preferencias, por una teoría sobre otra, 60, 101, 111, 113, 117.
Preguntas de qué-es, 176, 301-06 y *n*. *Véase también:* Definiciones, doctrina esencialista de. Esencia. Esencialismo.
Prejuicios, 87-88, 196-97, 294-301.
Presuposiciones, *véase*: Suposiciones.
Principio de economía de Mach, 374.
Probabilidad, probabilidades, teoría de, Parte II, 161*n*, 257-63, 267-69, 282*tn*, 292, 293*n*, 303, 327, 343, 345-46, 348-49, 355 y *n*, 361-62, 364-67, 375*n*, 382-85, 392, 395, 398, 399*t*, 405, 407, 410-11, 412, 413*t*, 414*t*, 418 y *n*, 420*t*, 421*t*, 422*t*, 424*t*, 425-29, 433*n*, 434*t*, 438-40.
  absoluta *versus* relativa, 287, 323-27, 349, 353, 366-67, 404*n*.
  afectada por la información, 433*n*.
  a priori, 324.
  Carnap sobre, 340, 377*n*.
  como generalización de la lógica deductiva, 380.
  como grado de creencia racional, 427.
  como límite de la frecuencia, 418.
  como medida de las posibilidades, 323, 326, 397, 413, 426, 434.
  como medio de juzgar a las hipótesis, 46, 59, 263, 283, 292*n*.
  como propensión, 300.
  como propiedad de las condiciones generadoras, 395.
  como propiedad de una sucesión, 395.
  como propiedad disposicional, 398.
  como sustituto de la certidumbre, 265.
  comparada con el contenido, 264, 295, 324.
  comparada con la corroboración, 263-67, 271, 287-91, 294.
  de hipótesis, en oposición a -de un suceso, 264-67, 332, 352, 382, 384-86, 418, 424.
  de inferencia, 267.
  de obtener ciertos resultados experimentales, 330, 339-40.
  definición clásica de, 414.
  definiciones de, 420-23.
  dependencia de las condiciones objetivas del experimento, 389.
  estadística, 376.
  estimaciones a partir de la experiencia pasada, 342-43.
  explicación inductivista de, 260, 368*t*, 373, 376, 377-78, 380.
  grado de, 189, 257, 269, 272, 295.
  iguales, 414-15, 418, 428, 431-32.
  interpretación frecuencial de, 324.
  numéricas, 327, 416.
  recurso inductivista a, 259.
  Russell sobre, 107.
  sentidos de, 322-24, 341-42.
  singular, 333, 335, 387, 389, 390.

394-95, 398, 404*n*, 413, 424, 430, 438.
supuestamente ampliativa, 41–42.
Véase también: Corroboración, grados de. Verosimilitud.
Problemas, 26, 29, 45, 48, 64-66, 70, 78, 82, 83, 89, 90, 95, 197, 200, 207, 216, 234, 315-16.
Procedimientos inductivos, Parte I, sección 3, 74-79, 87, 90, 158, 162, 295, 299, 373.
Profundidad, grados de, 179 y *n*.
condiciones suficientes de, 179, 185 y *n*.
del mundo, 145, 299.
de la ciencia, 298.
de una teoría, 178, 184, 190.
idea de, 139.
problema de, 181-85.
Véase también: Contenido. Simplicidad.
Programas.
de Carnap, 373-74.
metafísicos de investigación, 31*n*, 232*n*, 333.
Programas de investigación, 31*n*, 34*n*, 400; *véase también:* Metafísica, programas de investigación de.
Propensión, propensiones, 322*n*, 326-28, 331, 352, 377, 391, 396-400, 414, 430, 434, 437-38.
como disposicionales, 391, 398, 436.
como físicamente reales, 322*n*, 323, 326, 387.
como perfeccionamiento de las concepciones de Peirce, 322*n*.
como ponderaciones de las posibilidades, 390, 397.
como propiedades relacionales de la situación objetiva total, 398, 434.
como tendencias físicas a realizar posibilidades, 326, 329.
contrastadas por medio de las distribuciones de las frecuencias, 330, 437.
diferentes de las potencialidades aristotélicas, 398.
explican y predicen propiedades estadísticas de sucesiones, 390.
idea de, atribuida erróneamente a Peirce, 322*n*.
objetivas, 389.
supuestamente, una concepción antropomórfica, 396, 398.
teoría del autor de, 322*n*, 388-400.
y campos de fuerzas, 326, 391, 398, 437.
y enunciados singulares, 326, 327, 438.
Véase también: Cálculo de probabilidades, interpretación propensivista.
Prueba, pruebas, 26, 35, 310, 364.
Prueba y error, método de, 75, 79-82, 91-92, 116, 138, 220.
Pseudo-ciencia, 23, 203, 214, 215, 229-33.
Pseudo-explicaciones, 396.
Pseudo-problema, 200, 216.
Pseudo-proposiciones, 15, 16 y *n*, 158.
Psicoanálisis, Parte I, sección 18, 29, 202, 208*n*, 209, 216 y *n*, 232.
Psicología, 24, 131, 147, 194, 202, 232, 304*n*, 374.
Puente, puentes.
crítica de von Mises de, 419.
de hipótesis probabilísticas a contrastaciones de frecuencias relativas, 418.
entre enunciados frecuenciales y probabilísticos, 418, 421.
entre la teoría clásica y la frecuencial, 437.
entre la teoría clásica y la estadística, 420.
entre la teoría frecuencial y la de la medida, 387, 400.
problema de, 418.
y teoría propensivista, 421.
Véase también: Correspondencia.

Racionalidad, 67*t*, 91, 97.
Racionalismo 27, 46*t*, 47, 59, 67-68, 72, 196, 230*n*.
Radioactividad, 81.
Rayos X, 81.
Razón, razones, razonamiento, Parte I, sección 2.
críticas (opuesto a positivas o justificativas), 60, 63-69, 72, 195.
deductiva, 77, 364.
diferentes de causas, 141.
humana, 163, 300.
inductiva, Parte I, capítulo I, 77, 364.
intento de Hume de mostrar la falta de validez de, 71.
positiva, 59-62, 65, 68, 70, 73.
y realidad, intento de Hegel de identificar, 78.
Reactividad, 139*t*.
Realimentación, 39.
Realidad, 29, 121, 124, 147, 169, 176*n*, 313, 326, 341, 391.

de los cuerpos físicos, 169, 194.
Realismo, Parte I, secciones 5-16, 97, 128, 131, 142, 152, 171, 172, 186, 197.
de Carnap, 167.
de Hume, 163.
lógicamente más fuerte que el idealismo, 143.
metafísico, parte I, sección 7, 29, 51, 120-22, 127-28, 155, 166, 168, 186, 189-98.
no hay justificación de, 125.
Reichenbach sobre, 163, 166-67.
Russel sobre, 126, 163.
verdad de, no establecida por la debilidad del idealismo de Hume, 151.
y el objetivo de la ciencia, 171, Parte I, sección 15.
Recuento de posibilidades favorables, criticado por Hosiasson y Jeffreys, 415, 417.
Reflejo, 84-85*n*, 140.
Relexividad, ley de, 259*n*.
Refutación, *véase:* Falsación.
Regla inductiva simple, 333 y *n*, 341-47 y *n*, 351-52, 354, 357-58, 361 y *n*, 362-65, 368*t*, 369*n*, 370, 371, 375-76, 380. *Véase también:* Falacia del jugador.
Regularidades, 75, 97, 112, 116-17, 141, 412, 438. *Véase también:* Leyes de la naturaleza.
Relatividad, teoría de, 31-32, 65, 66*n*, 111, 287.
Relativismo, 25, 58, 64, 66, 196.
Religión, 24, 146, 163, 208.
Repetibilidad, 336.
Repetición, 82, 91, 258, 350, 363.
aprendizaje por, 74, 79-80, 82-85, 138-40.
de observaciones, 72, 75, 138, 139, 318.
de un experimento, 329-31, 337--39, 344, 352, 358, 398, 430.
sucesiones de, 390, 414.
Véase también: Independencia.
Resolución de problemas, 311*n*, 312, 316.
Resonancia, teoría de, 191*t*.
Resultados experimentales, 90, 328-31, 357-58, 409, 427, 436.
delimitación, 427.
efecto en resultado de otros resultados posteriores, 338.
fiabilidad, 87.
observación de, 336.
probabilidad de, 330, 339.
Véase también: Efectos secundarios. Condiciones experimentales. Experimento. Independencia.
Retrogresión infinita, en el argumento justificacionista, 69, 371.
Revolución, revoluciones, 35.
científicas, 27-28, 35.
estructura lógica de la, 36.
ideológicas, 35.
identificación de Kuhn de ambas, 35.
separación del autor de las ideológicas de las científicas, 35.
Rigidez, 178, 179*n*.
Russell, cuestión fundamental de, 126*t*-27, 132, 134, 381.

Saber, república del, 300.
Satisfacción de deseos, 204-13 y *n*.
Seguros, 350-52.
Selección, en probabilidades, 388*n*, 401-02*n*, 403, 405, 412-13*n*, 425.
Sensacionalismo, 125.
Sensibilidad, 139*t*.
Sentido (significado), 149, 237, 250, 305, 307, 422.
análisis, 216, 301-17, 399.
comparado con contrastabilidad, 214-18, 249.
criterio de, 158, 168, 214 y *n*, 215-18, 234-35 y *n*, 255*n*, 256.
problema de la carencia de, 216.
teoría observacionista o fenomenalista de, 148.
teoría operacionalista de, 250.
y enfoque inductivista, 250, 251, 254.
Véase también: Filosofía, programas de Wittgenstein para.
Sentido común, 55, 95, 142, 169, 170*n*.
ciencia como, 299.
choque con, 171.
enunciados de, y realidad de los cuerpos físicos y existencia de las leyes de la naturaleza, 169.
no influenciado por la filosofía y la teología, 126.
y concepción clásica empirista, 86.
Símbolos, simbolismo, 148, 236, 242, 253, 376.
Simplicidad, 47, 179, 265, 302, 418.
Simultaneidad, 303*n*.
Sistema solar, 98, 101.
Sistemas de juego, 345, 369, 403, 411, 425, 438.
calculabilidad y, 426, 433.

de éxito, 430.
insensibilidad respecto a, 402, 420.
ley de la exclusión de los, 339-340, 403, 411, 413, 418, 421, 425.
Véase también: Aleatoriedad, axioma de von Mises de.
Situación de los problemas, 33, 70, 90, 313, 337.
Solipsismo, 123-24, 381. *Véase también:* Subjetivismo.
Subjetivismo, 24, 121, 133-35, 140-41, 143, 336-40, 341, 347, 352-53, 355-57, 380-81, 391.
Sucesión, sucesiones, de sucesos o contrastaciones, 196*n*, 327, 345*n*, 369, 390-06, 407-10, 414, 418*n*, 421-23, 425-27, 432.
aleatorias, 369, 388, 401, 403 y *n*, 404, 422, 436.
alternativas, 425.
bernoullianas, 402*n*.
caracterizadas por las condiciones generadoras, 395, 397-99.
convergentes y divergentes, 405-06, 411.
de referencia, 394, 422, 423.
dependientes, 344, 345, 346, 370*n*.
estabilizados, 370.
finitas e infinitas, 389, 391-94, 406, 409, 419, 425, 429, 438.
ilimitadas, 388.
límite de, 307, 309.
límite de la frecuencia relativa de, 405-10.
normales 402*nt*, 421*n*.
posibles, 424*t*.
reales, 327, 395, 398, 424.
suceso singular, probabilidad dentro de, 389-90.
virtuales, 327, 395.
Véase también: Sucesiones alternativas. Colectivos. Números admisibles. Cálculo de probabilidades, interpretaciones de.
Sucesiones alternativas, 404*t* y *n*, 406, 408-09, 424 y *n*.
formación de colectivo a partir de, 408.
resultados de Borel extendidos a, 404.
subconjuntos de, 424.
y frecuencias convergentes, 408.
y límites de la frecuencia, 407, 423*n*.
Sucesiones fortuitas aleatoriformes, teoría matemática de, 404.
Suceso, sucesos, 332, 352, 418-23, 424, 427-28, 431.
aleatorios, 419.
condiciones objetivas de, 389.
complejos, 409.
empíricos, 419.
fortuitos, 410, 412.
frecuencia de, 410.
improbables, 418*n*.
independientes, 410, 438 y ss.
observables, 276-83, 285, 309.
probabilidad de, *véase:* Probabilidad, de un suceso.
probabilidad *versus* frecuencia de, 421.
secuencias infinitas de, 438.
singulares, 327-29, 360, 362, 389, 395, 398, 413, 427, 429, 431, 438.
interpretación frecuencial de, 327, 389.
interpretación propensivista de, 426-27.
probabilidades inherentes de, 396.
totalidad de, 118.
Sueños, Parte I, sección, 18, 206, 212*n*.
análisis de Freud de, 204-13.
complacientes, 209 y *n*, 210.
contenido latente y manifiesto de, 205, 207 y *n*.
de angustia, 205-13.
y hambre, 212, 213*n*.
Sugestión, inconsciente, 211.
Suma, ley general de la, 259*n*, 427.
Supuestos.
básicos, conscientes e inconscientes, 54.
mantenidos en común, o ambientales, 42, 46, 57.
metodológicos, 114.
no demostrables, injustificables, 68.
psicológicos, 115.
teóricos, 250.
«últimos», no son necesarios, 69.
Sustancia, 144, 145 y *n*.

Tabula rasa, 139.
Tautología, 253, 271, 283, 285, 288, 324, 333, 349, 367, 377 y *n*, 379, 431; *véase también:* Enunciados analíticos.
Tendencias, *véase:* Propensiones.
Teología, y ataque positivista, 252.
Teorema de Poisson, 419.
Teoría, teorías, sistemas teóricos, *véase:* Expectativa. Hipótesis. Leyes de la naturlaeza.
Teoría atómica, 31.
cuestión de su significatividad, 145.

fundamento de, 30.
relación con la ciencia y la filosofía, 145.
Teoría copernicana.
e instrumentalismo, 156.
y Bacon, 148.
Teoría de la medida, 420, 427.
Teoría de Misner-Wheeler, 66.
Teoría freudiana, Parte 1, sección 18.
Teoría, teorías de campos.
de Einstein y Weyl, 31.
de Faraday y Maxwell, 31.
Teorías de las mareas, 109, 229, 230*n*.
Términos, teóricos, 249-54, 263, 313-14.
Termodinámica, 200.
Tiempo, 117-18.
Tradición, 68, 132-35, 171, 230.
Transcendentalismo, *véase*: Argumento, transcendental. Crítica transcendental.

Universales, 109, 150.
Universalidad, grado de y niveles de 112, 174, 176-177, 185.
Universo, 129-30, 167, 384.
Utilidad decreciente, ley de, 356, 369.

Verdad, 39-40, 69, 299.
a priori, 193.
aproximación a la, 39-40, 46, 53, 63-67, 96-99, 100, 107, 170, 193, 197, 341, 386.
como norma de crítica, 65-66.
concepciones de Kuhn sobre, 35.
concepto de Tarski sobre, 113, 119, 305, 312-14.
criterio de, 35.
explicación del propio autor, 35, 39-40.
Polanyi sobre, 35.
problema de, 23.
posición relativista, 64.
teoría de la correspondencia, 23, 35, 66.
Véase también: Verosimilitud.
Verificabilidad, 214-15, 222-23, 234, 239.
Verificación, 46, 170, 202-04, 208-10, 214-15, 221-29, 262, 274, 296, 298, 303*n*, 372.
Verificacionismo, 38, 203-14, 222, 269, 273-76.
Verosimilitud, 39-40, 97-98*n*, 101-02, 107*n*.

En este primer volumen del *Post Scriptum*, obra publicada tras más de veinticinco años de haber sido escrita, Popper formula y explica su teoría no justificacionista del conocimiento. La ciencia —ciencia empírica— busca teorías explicativas *verdaderas;* pero nunca puede demostrar o establecer definitivamente o justificar sus teorías como verdaderas, ni siquiera si son, en efecto, verdaderas. Así pues, la ciencia tiene que continuar cuestionando y criticando todas sus teorías, incluso las que son verdaderas. Sin embargo, puede dar razones para preferir unas teorías a otras: algunas están mejor contrastadas (corroboradas), pueden plantear o resolver más problemas, tener mayor poder explicativo e incluso ser capaces de corregir aquellos hechos (aparentes) que debían «explicar». De este modo, aunque no un criterio para conocer la verdad, sí poseemos algo parecido a un criterio sobre el progreso de nuestras teorías: del progreso hacia la verdad y hacia explicaciones teóricas más profundas y eficaces. Popper sustituye el objetivo utópico de «justificar» nuestras teorías por el objetivo realista de someterlas a la crítica empírica más severa que nuestro ingenio pueda idear. Contrastar una teoría no consiste sólo en observar, sino en realizar intentos ingeniosos de encontrar sus puntos débiles, los puntos en los que la teoría puede ser deficiente, los problemas que puede ser incapaz de resolver.

Sir Karl Popper es doctor «honoris causa» por catorce universidades estadounidenses, británicas, alemanas, austríacas, neozelandesas y canadienses. Es miembro de número u honorario de doce Academias. Algunas de sus obras se han traducido a veintidós idiomas.

La presente edición está preparada por William Warren Bartley III, graduado por Harvard y por la Universidad de Londres. Profesor en las universidades de Londres, Cambridge y Pittsburgh, actualmente es catedrático de Filosofía en la Universidad del Estado de California, en Hayward.

TIT: REALISMO Y EL OBJETIVO DE LA CIENCIA
AUTOR: POPPER, K.
CODIGO: TE7401228 PRECIO: 2,805.00

Filosofía y Ensayo

tecnos